(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 August 2003 (14.08.2003)

PCT

(10) International Publication Number
WO 03/066675 A1

(51) International Patent Classification[7]: C07K 14/415, C12N 15/82, C07K 16/16, G01N 33/50, C12N 5/10, A01H 1/04

(21) International Application Number: PCT/NL03/00091

(22) International Filing Date: 7 February 2003 (07.02.2003)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
02075565.8 8 February 2002 (08.02.2002) EP

(71) Applicant *(for all designated States except US)*: **KWEEKEN RESEARCHBEDRIJF AGRICO B.V.** [NL/NL]; Burchtweg 17, NL-8314 PP Bant (NL).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **ALLEFS, Josephus, Jacobus, Hendricus, Maria** [NL/NL]; Naardenstraat 8, NL-8304 EJ Emmeloord (NL). **VAN DER VOSSEN, Edwin, Andries, Gerard** [NL/NL]; Oostlaan 20, NL-3572 ZM Utrecht (NL).

(74) Agent: **PRINS, A.W.**; Nieuwe Parklaan 97, NL-2587 BN Den Haag (NL).

(81) Designated States *(national)*: AE, AG, AL, AM, AT (utility model), AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ (utility model), CZ, DE (utility model), DE, DK (utility model), DK, DM, DZ, EC, EE (utility model), EE, ES, FI (utility model), FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK (utility model), SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:
-- *with international search report*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 03/066675 A1

(54) Title: GENE CONFERRING RESISTANCE TO PHYTOPHTHORA INFESTANS (LATE-BLIGHT) IN SOLANACEA

(57) Abstract: The invention relates to the field of plant diseases, in particular to oomycete infections such as late blight, a disease of major importance to production of *Solanaceae* such as potato and tomato cultivars. The invention provides a method for providing a plant or its progeny with resistance against an oomycete infection comprising providing said plant or part thereof with a gene or functional fragment thereof comprising a nucleic acid, said nucleic acid encoding a gene product that is capable of providing a member of the *Solanaceae* with resistance against an oomycete fungus.

GENE CONFERRING RESISTANCE TO PHYTO§§THERA INFESTANS (LATE-BLIGHT) IN SOLANACEAE

Late blight, caused by the oomycete pathogen *Phytophthora infestans* is world-wide the most destructive disease for potato cultivation. The disease also threatens the tomato crop. The urgency of obtaining resistant cultivars has intensified as more virulent, crop-specialised and pesticide resistant strains of the pathogen are rapidly emerging.

A way to prevent crop failures or reduced yields is the application of fungicides that prevent or cure an infection by *P. infestans*. However, the application of crop protectants is widely considered to be a burden for the environment. Thus, in several Western countries, legislation is becoming more restrictive and partly prohibitive to the application of specific fungicides, making chemical control of the disease more difficult. An alternative approach is the use of cultivars that harbour partial or complete resistance to late blight. Two types of resistance to late blight have been described and used in potato breeding. One kind is conferred by a series of major, dominant genes that render the host incompatible with specific races of the pathogen (race specific resistance). Eleven such *R* genes (*R1-R11*) have been identified and are believed to have originated in the wild potato species *Solanum demissum*, which is native to Mexico, where the greatest genetic variation of the pathogen is found. Several of these *R* genes have been mapped on the genetic map of potato (reviewed in Gebhardt and Valkonen, 2001 Annu. Rev. Phytopathol. 39: 79-102). *R1* and *R2* are located on chromosomes 5 and 4, respectively. *R3, R6* and *R7* are located on chromosome 11. Unknown *R* genes conferring race specific resistance to late blight have also been described in *S. tuberosum* ssp. *andigena* and *S. berthaultii* (Ewing et al., 2000 Mol. Breeding 6: 25-36). Because of the high level of resistance and ease of transfer, many cultivars contain *S. demissum* derived resistance. Unfortunately, *S. demissum* derived race specific resistance, although nearly complete, is not durable. Once newly bred cultivars are grown on larger scale in commercial fields, new virulences emerge in *P. infestans* that render the pathogen able to overcome the introgressed resistance. The second type of resistance, termed field resistance and often quantitative in nature, is thought to be race non-specific and more durable. Field resistance to late blight can be found in several Mexican and Middle and South American *Solanum* species (Rossi et al., 1986 PNAS 95:9750-9754).

Diploid *S. bulbocastanum* from Mexico and Guatemala is one of the tuber bearing species that is known for its high levels of field resistance to late blight (Niederhauser and Mills, 1953 Phytopathology 43: 456-457). Despite differences in endosperm balance numbers, introgression of the *S. bulbocastanum* resistance trait has been successful. Ploidy manipulations and a series of tedious bridge crosses has resulted in *S. bulbocastanum* derived, *P. infestans* resistant germplasm (Hermsen and Ramanna, 1969 Euphytica 18:27-35; 1973 Euphytica 22:457-466; Ramanna and Hermsen, 1971 Euphytica 20:470-481; Hermsen and De Boer, 1971 Euphytica 20:171-180). However, almost 40 years after the first crosses and intense and continuous breeding efforts by potato breeders in the Netherlands with this germplasm, late blight resistant cultivars still remain to be introduced on the market. Successful production of somatic hybrids of *S. bulbocastanum* and *S. tuberosum* has also been reported (Thieme et al., 1997 Euphytica 97(2):189-200; Helgeson et al., 1998 Theor Appl. Genet 96:738-742). Some of these hybrids and backcrossed germplasm were found to be highly resistant to late blight, even under extreme disease pressure. Despite reports of suppression of recombination, resistance in the backcrossed material appeared to be on chromosome 8 within an approximately 6 cM interval between the RFLP markers CP53 and CT64 (Naess et al., 2000 Theor. Appl Genet 101:697-704). A CAPS marker derived from the tomato RFLP probe CT88 cosegregated with resistance. Suppression of recombination between the *S. bulbocastanum* and *S. tuberosum* chromosomes forms a potential obstacle for successful reconstitution of the recurrent cultivated potato germplasm to a level that could meet the standards for newly bred potato cultivars. Isolation of the genes that code for resistance found in *S. bulbocastanum* and subsequent transformation of existing cultivars with these genes, would be a much more straight forward and quicker approach when compared to introgression breeding.

The cloning and molecular characterisation of numerous plant *R* genes conferring disease resistance to bacteria, fungi, viruses, nematodes, and insects has identified several structural features characteristic to plant *R* genes (reviewed in Dangl and Jones, 2001 Nature 411, 826-833). The majority are members of tightly linked multigene families and all *R* genes characterised so far, with the exception of Pto, encode leucine-rich repeats (LRRs), structures shown to be involved in protein-protein interactions. LRR containing *R* genes can be divided into two classes based on the presence of a putative tripartite nucleotide-binding site (NBS). *R* genes of the NBS-LRR class comprise motifs that are shared with animal apoptosis regulatory

proteins (van der Biezen et al., 1998 Curr. Biol. 8, 226-227; Aravind et al., 1999 Trends Biochem. Sci. 24, 47-53) and can be subdivided into two subgroups based on their N-terminal domain, which either exhibits sequence similarity to the Drosophila Toll protein and the mammalian interleukin-1 receptor domain (TIR-NBS-LRR), or contains a potential leucine zipper or coiled-coil domain (CC-NBS-LRR; Pan et al., 2000 Genetics. 155:309-22). LRR *R* genes without an NBS encode transmembrane proteins, whose extracellular N-terminal region is composed of LRRs (Jones et al., 1994 Adv. Bot. Res.24, 89-167). These genes can be divided into two subgroups based on the presence of a cytosolic serine/threonine kinase domain (Song et al., 1995 Science, 270, 1804-1806). Four *R* genes have currently been cloned from potato. All four, including the *S. demissum* derived *R1* gene conferring race specific resistance to late blight, belong to the CC-NBS-LRR class of plant *R* genes (Bendahmane et al., 1999 Plant Cell 11, 781-791; Bendahmane et al., 2000 Plant J. 21, 73-81; van der Vossen et al., 2000 Plant Journal 23, 567-576; Ballvora et al., 2002 Plant Journal 30, 361-371).

The invention provides an isolated or recombinant nucleic acid comprising a nucleic acid coding for the amino acid sequence of fig. 8 or a functional fragment or a homologue thereof. The protein coded by said amino acid has been detected as being member of a cluster of genes identifiable by phylogenetic tree analysis, which thus far consists of the proteins *Rpi-blb, RGC1-blb, RGC3-blb and RGC4-blb* (herein also called the *Rpi-blb* gene cluster) of figure 9.

Phylogenetic tree analysis is carried out as follows. First a multiple sequence alignment is made of the nucleic acid sequences and/or preferably of the deduced amino acid sequences of the genes to be analysed using CLUSTALW (http://www2.ebi.ac.uk/clustalw), which is in standard use in the art. ClustalW produces a .dnd file, which can be read by TREEVIEW (http://taxonomy.zoology.gla.ac.uk/rod/rod.html). The phylogenetic tree depicted in Figure 9A is a phylogram.

Phylogenetic studies of the deduced amino acid sequences of *Rpi-blb, RGC1-blb, RGC3-blb, RGC4-blb* and those of the most similar genes from the art (as defined by the BLASTX) derived from diverse species, using the Neighbour-Joining method of Saitou and Nei (1987 Molecular Biology and Evolution 4, 406-425), shows that corresponding genes or functional fragments thereof of the *Rpi-blb* gene cluster can be placed in a separate branch (Figure 9A).

Sequence comparisons between the four members of the *Rpi-blb* gene cluster identified on 8005-8 BAC clone SPB4 show that sequence homology within the *Rpi-blb* gene cluster varies between 70% and 81% at the amino acid sequence level. The deduced amino acid sequence of *Rpi-blb* shares the highest overall homology with *RGC3-blb* (81% amino-acid sequence identity; Table 4). When the different domains are compared it is clear that the effector domains present in the N-terminal halves of the proteins (coiled-coil and NBS-ARC domains) share a higher degree of homology (91% sequence identity) than the C-terminal halves of these proteins which are thought to contain the recognition domains (LRRs; 71% amino acid sequence identity). Comparison of all four amino-acid sequences revealed a total of 104 Rpi-blb specific amino acid residues (Figure 10). The majority of these are located in the LRR region (80/104). Within the latter region, these specific residues are concentrated in the LRR subdomain xxLxLxxxx. The relative frequency of these specific amino-acid residues within this LRR subdomain is more than two times higher (28.3%) than that observed in the rest of the LRR domain (12.3%). The residues positioned around the two conserved leucine residues in the consensus xxLxxLxxxx are thought to be solvent exposed and are therefore likely to be involved in creating/maintaining recognition specificity of the resistance protein.

Sequences of additional members of the *Rpi-blb* gene cluster can be obtained by screening genomic DNA or insert libraries, e.g. BAC libraries with primers based on signature sequences of the *Rpi-blb* gene. Screening of various *Solanum* BAC libraries with primer sets A and/or B (Table 2 and Figure 7) identified numerous *Rpi-blb* homologues derived from different *Solanum* species. Alignment of these additional sequences with those presented in Figure 10 will help identify additional members of the *Rpi-blb* gene cluster and specific amino acid residues therein responsible for *P. infestans* resistance specificity. Furthermore, testing additional sequences in the above described phylogenetic tree analyses, e.g. using the Neighbour-Joining method of Saitou and Nei (1987 Molecular Biology and Evolution 4, 406-425), provides additional identification of genes belonging to the *Rpi-blb* gene cluster.

The invention provides the development of an intraspecific mapping population of *S. bulbocastanum* that segregated for race non-specific resistance to late blight. The resistance was mapped on chromosome 8, in a region located 0.3 cM distal from CT88. Due to the race non-specific nature of the resistance, *S. bulbocastanum* late blight resistance has always been thought to be *R* gene

independent. However, with the current invention we demonstrate for the first time that *S. bulbocastanum* race non-specific resistance is in fact conferred by a gene bearing similarity to an *R* gene of the NBS-LRR type.

The invention further provides the molecular analysis of this genomic region and the isolation by map based cloning of a DNA-fragment of the resistant parent that harbours an *R* gene, designated *Rpi-blb*. This DNA-fragment was subcloned from an approximately 80 kb bacterial artificial chromosome (BAC) clone which contained four complete *R* gene-like sequences in a cluster-like arrangement. Transformation of a susceptible potato cultivar by *Agrobacterium tumefaciens* revealed that one of the four *R* gene-like sequences corresponds to *Rpi-blb* that provides the race non-specific resistance to late blight. Characterisation of the *Rpi-blb* gene showed that it is a member of the NBS-LRR class of plant *R* genes. The closest functionally characterised sequences of the prior art are members of the *I2* resistance gene family in tomato. These sequences have an overall amino acid sequence identity of approximately 32% with that of *Rpi-blb*.

Thus, in a first embodiment, the invention provides an isolated or recombinant nucleic acid, said nucleic acid encoding a gene product having the sequence of *Rpi*-blb or a functional fragment thereof that is capable of providing a member of the *Solanaceae* family with race non-specific resistance against an oomycete pathogen.

Isolation of the gene as provided here that codes for the desired resistance trait against late blight and subsequent transformation of existing potato and tomato cultivars with this gene now provides a much more straightforward and quicker approach when compared to introgression breeding. The results provided here offer possibilities to further study the molecular basis of the plant pathogen interaction, the ecological role of *R* genes in a wild Mexican potato species and are useful for development of resistant potato or tomato cultivars by means of genetic modification.

In contrast to the *R* genes cloned and described so far, the gene we provide here is the first isolated *R* gene from a *Solanum* species that provides race non-specific resistance against an oomycete pathogen. Notably, the invention provides here a nucleic acid wherein said *Solanum* species that is provided with the desired resistance comprises *S. tuberosum*. In particular, it is the first gene that has been isolated from a phylogenetically distinct relative of cultivated potato, *S. bulbocastanum*, for which it was shown by complementation assays, that it is functional in *S. tuberosum*. These data imply that the gene *Rpi-blb* can easily be

applied in potato production without a need for time-consuming and complex introgression breeding.

The following definitions are provided for terms used in the description and examples that follow.

- *Nucleic acid*: a double or single stranded DNA or RNA molecule.
- *Oligonucleotide*: a short single-stranded nucleic acid molecule.
- *Primer*: the term primer refers to an oligonucleotide that can prime the synthesis of nucleic acid.
- *Homology*: homology is the term used for the similarity or identity of biological sequence information. Homology may be found at the nucleotide sequence and/or encoded amino acid sequence level. For calculation of precentage identity the BLAST algorithm can be used (Altschul *et al.*, 1997 Nucl. Acids Res. 25:3389-3402) using default parameters or, alternatively, the GAP algorithm (Needleman and Wunsch, 1970 J. Mol. Biol. 48:443-453), using default parameters, which both are included in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, Wisconsin, USA. BLAST searches assume that proteins can be modelled as random sequences. However, many real proteins comprise regions of nonrandom sequences which may be homopolymeric tracts, short-period repeats, or regions enriched in one or more amino acids. Such low-complexity regions may be aligned between unrelated proteins even though other regions of the protein are entirely dissimilar. A number of low-complexity filter programs can be employed to reduce such low-complexity alignments. For example, the SEG (Wooten and Federhen, 1993 Comput. Chem. 17:149-163) and XNU (Claverie and States, 1993 Comput. Chem. 17:191-201) low-complexity filters can be employed alone or in combination.

  As used herein, 'sequence identity' or 'identity' in the context of two protein sequences (or nucleotide sequences) includes reference to the residues in the two sequences which are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognised that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acids are substituted for other amino acid residues with similar chemical properties (e.g. charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percentage

sequence identity may be adjusted upwards to correct for the conservative nature of the substitutions. Sequences, which differ by such conservative substitutions are said to have 'sequence similarity' or 'similarity'. Means for making these adjustments are well known to persons skilled in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is give a score of zero, a conservative substitution is given a score between 0 and 1. The scoring of conservative substitutions is calculated, e.g. according to the algorithm of Meyers and Miller (Computer Applic. Biol. Sci. 4:11-17, 1988).

As used herein, 'percentage of sequence identity' means the value determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the amino acid sequence or nucleotide sequence in the comparison window may comprise additions or deletions (i.e., gaps) as compared to the reference sequence for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical amino acid or nucleic acid base residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Preferably the amino acid sequence of the protein of the invention shares at least 82% or higher homology with the sequence as depicted in Fig. 8. As shown in Table 4, the closest functionally characterised sequence of the prior art (members of the *I2 Fusarium* resistance gene cluster in tomato) has a much lower level of amino acid sequence identity than this (32% with respect to that of *Rpi-blb*). Homology within the gene cluster of the present invention varies between 70% and 81% at the amino acid sequence level.

Homologous nucleic acid sequences are nucleic acid sequences coding for a homologous protein defined as above. One example of such a nucleic acid is the sequence as provided in figure 6A. However, there are many sequences which code for a protein which is 100% identical to the protein as depicted in fig. 8. This is due to the 'wobble' in the nucleotide triplets, where more than one triplet can code for one and the same amino acid. Thus, even without having an effect on the amino acid sequence of the protein the nucleotide sequence coding for this protein can be varied substantially. It is acknowledged that nucleotide sequences coding

for amino acid sequences that are not 100% identical to said protein can contain even more variations. Therefore, the percentage identity on nucleic acid sequence level can vary within wider limits, without departing from the invention.

- *Promoter*: the term "promoter" is intended to mean a short DNA sequence to which RNA polymerase and/or other transcription initiation factors bind prior to transcription of the DNA to which the promoter is functionally connected, allowing transcription to take place. The promoter is usually situated upstream (5') of the coding sequence. In its broader scope, the term "promoter" includes the RNA polymerase binding site as well as regulatory sequence elements located within several hundreds of base pairs, occasionally even further away, from the transcription start site. Such regulatory sequences are, e.g., sequences that are involved in the binding of protein factors that control the effectiveness of transcription initiation in response to physiological conditions. The promoter region should be functional in the host cell and preferably corresponds to the natural promoter region of the *Rpi-blb* resistance gene. However, any heterologous promoter region can be used as long as it is functional in the host cell where expression is desired. The heterologous promoter can be either constitutive or regulatable, tissue specific or not specific. A constitutive promoter such as the CaMV 35S promoter or T-DNA promoters, all well known to those skilled in the art, is a promoter which is subjected to substantially no regulation such as induction or repression, but which allows for a steady and substantially unchanged transcription of the DNA sequence to which it is functionally bound in all active cells of the organism provided that other requirements for the transcription to take place is fulfilled. It is possible to use a tissue-specific promoter, which is driving expression in those parts of the plant which are prone to pathogen infection. In the case of *Phytophthora* a promoter which drives expression in the leaves, such as the ferredoxin promoter, can be used. A regulatable promoter is a promoter of which the function is regulated by one or more factors. These factors may either be such which by their presence ensure expression of the relevant DNA sequence or may, alternatively, be such which suppress the expression of the DNA sequence so that their absence causes the DNA sequence to be expressed. Thus, the promoter and optionally its associated regulatory sequence may be activated by the presence or absence of one or more factors to affect transcription of the DNA sequences of the genetic construct of the

invention. Suitable promoter sequences and means for obtaining an increased transcription and expression are known to those skilled in the art.

- *Terminator*: the transcription terminator serves to terminate the transcription of the DNA into RNA and is preferably selected from the group consisting of plant transcription terminator sequences, bacterial transcription terminator sequences and plant virus terminator sequences known to those skilled in the art.
- *Gene*: the term "gene" is used to indicate a DNA sequence which is involved in producing a polypeptide chain and which includes regions preceding and following the coding region (5'-upstream and 3'-downstream sequences) as well as intervening sequences, the so-called introns, which are placed between individual coding segments (so-called exons) or in the 5'-upstream or 3'-downstream region. The 5'-upstream region may comprise a regulatory sequence that controls the expression of the gene, typically a promoter. The 3'-downstream region may comprise sequences which are involved in termination of transcription of the gene and optionally sequences responsible for polyadenylation of the transcript and the 3' untranslated region. The term "resistance gene" is an isolated nucleic acid according to the invention said nucleic acid encoding a gene product that is capable of providing a plant with resistance against a pathogen, more specifically said plant being a member of the *Solanaceae* family, more preferably potato or tomato, said pathogen more specifically being an oomycete pathogen, more specifically *Phytophthora*, more specifically *Phytophthora infestans*, said nucleic acid preferably comprising a sequence as depicted in Fig. 8 or part thereof, or a homologous sequence with essentially similar functional and structural characteristics. A functionally equivalent fragment of such a resistance gene or nucleic acid as provided by the invention encodes a fragment of a polypeptide having an amino acid sequence as depicted in Fig. 8 or part thereof, or a homologous and/or functionally equivalent polypeptide, said fragment exhibiting the characteristic of providing at least partial resistance to an oomycete infection such as caused by *P. infestans* when incorporated and expressed in a plant or plant cell.
- *Resistance gene product*: a polypeptide having an amino acid sequence as depicted in Fig. 8 or part thereof, or a homologous and/or functionally equivalent polypeptide exhibiting the characteristic of providing at least partial resistance to an oomycete infection such as caused by *P. infestans* when incorporated and expressed in a plant or plant cell.

Functionally equivalents of the protein of the invention are proteins that are homologous to and are obtained from the protein depicted in fig. 8 by replacing, adding and/or deleting one or more amino acids, while still retaining their pathogen resistance activity. Such equivalents can readily be made by protein engineering *in vivo*, e.g. by changing the open reading frame capable of encoding the protein so that the amino acid sequence is thereby affected. As long as the changes in the amino acid sequences do not altogether abolish the activity of the protein such equivalents are embraced in the present invention. Further, it should be understood that equivalents should be derivable from the protein depicted in fig. 8 while retaining biological activity, i.e. all, or a great part of the intermediates between the equivalent protein and the protein depicted in fig. 8 should have pathogen resistance activity. A great part would mean 30% or more of the intermediates, preferably 40% or more, more preferably 50% or more, more preferably 60% or more, more preferably 70% or more, more preferably 80% or more, more preferably 90% or more, more preferably 95% or more, more preferably 99% or more.

Preferred equivalents are equivalents in which the leucine rich repeat region is highly homologous to the LRR region as depicted in fig. 8. Other preferred equivalents are equivalents wherein the N-terminal effector domain is essential the same as the effector domain of *Rpi-blb*.

The protein of the invention comprises a distinct N-terminal effector domain and a leucine rich repeat domain. It is believed that conservation of these regions is essential for the function of the protein, although some variation is allowable. However, the other parts of the protein are less important for the function and may be more susceptible to change.

In order to provide a quick and simple test if the modified proteins and/or the gene constructs capable of expressing said modified proteins which are described here or any new constructs which are obvious to the person skilled in the art after reading this application indeed can yield a resistance response the person skilled in the art can perform a rapid transient expression test known under the name of ATTA (*Agrobacterium tumefaciens* Transient expression Assay). In this assay (of which a detailed description can be found in Van den Ackerveken, G., et al., Cell 87, 1307-1316, 1996) the nucleotide sequence coding for the modified protein which is to be tested is placed under

control of the CaMV 35S promoter and introduced into an *Agrobacterium* strain which is also used in protocols for stable transformation. After incubation of the bacteria with acetosyringon or any other phenolic compound which is known to enhance *Agrobacterium* T-DNA transfer, 1 ml of the Agrobacterium culture is infiltrated into an *in situ* plant leaf (from e.g. a tobacco or potato or tomato plant) by injection after which the plants are placed in a greenhouse and infected with a pathogen, preferably *P. infestans*. After 2-5 days the leaves can be scored for occurrence of resistance symptoms.

In the present invention we have identified and isolated the resistance gene *Rpi-blb*, which confers race non-specific resistance to *Phytophthora infestans*. The gene was cloned from a *Solanum bulbocastanum* genotype that is resistant to *P. infestans*. The isolated resistance gene according to the invention can be transferred to a susceptible host plant using *Agrobacterium* mediated transformation or any other known transformation method, and is involved in conferring the host plant resistant to plant pathogens, especially *P. infestans*. The host plant can be potato, tomato or any other plant, in particular a member of the *Solanaceae* family that may be infected by such a plant pathogen. The present invention provides also a nucleic acid sequence coding for this protein or a functional equivalent thereof, preferably comprising the *Rpi-blb* gene, which is depicted in Figure 6.

With the *Rpi-blb* resistance protein or functionally equivalent fragment thereof according to the invention, one has an effective means of control against plant pathogens, since the gene coding for the protein can be used for transforming susceptible plant genotypes thereby producing genetically transformed plants having a reduced susceptibility or being preferably resistant to a plant pathogen. In particular, a plant genetically transformed with the *Rpi-blb* resistance gene according to the invention has a reduced susceptibility to *P. infestans*.

In a preferred embodiment the *Rpi-blb* resistance gene comprises the coding sequence provided in Figure 6A or any homologous sequence or part thereof preceded by a promoter region and/or followed by a terminator region. The promoter region should be functional in plant cells, and preferably correspond to the native promoter region of the *Rpi-blb* gene. However, a heterologous promoter region that is functional in plant cells can be used in conjunction with the coding sequences.

In addition the invention relates to the *Rpi-blb* resistance protein which is encoded by the *Rpi-blb* gene according to the invention and which has an amino acid sequence provided in Figure 8, or a functional equivalent thereof.

The signal that triggers the expression of the resistance gene in the wild-type *S. bulbocastanum* or in the transgenic plants of the invention is probably caused by the presence of a pathogen, more specifically the pathogen *P. infestans*. Such systems are known for other pathogen-plant interactions (Klement, Z., In: Phytopathogenic Prokaryotes, Vol. 2, eds.: Mount, M.S. and Lacy, G.H., New York, Academic Press, 1982, pp. 149-177), and use of this system can be made to increase the applicability of the resistance protein resulting in a resistance to more pathogens (see EP 474 857). This system makes use of the elicitor compound derived from the pathogen and the corresponding resistance gene, wherein the resistance gene when activated by the presence of the elicitor would lead to local cell death (hypersensitive reaction). In case of the present resistance gene, the corresponding elicitor component has not yet been disclosed, but it is believed that this is achievable by a person skilled in the art. Once the elicitor component is isolated it will be possible to transform the gene coding for said elicitor together with the gene coding for the resistance protein into plant, whereby one of the genes is under control of a pathogen-inducible promoter. These promoters are well known in the art (e.g. *prp1, Fis1, Bet v 1, Vst1, gstA1,* and sesquiterpene cyclase, but any pathogen-inducible promoter which is switched on after pathogen infection can be used). If the transgenic plant contains such a system, then pathogen attack which is able to trigger the pathogen-inducible promoter will cause production of the component which is under control of said promoter, and this, in connection with the other component being expressed constitutively, will cause the resistance reaction to occur.

It will also be possible to mutate the resistance protein causing it to be in an active state (see EP1060257). Since this would permanently result in the resistance reaction to occur, which ultimately leads to local cell death, care should be taken not to constitutively express the resistance protein. This can be accomplished by placing the mutated resistance protein under control of a pathogen-inducible promoter, which not only would allow for expression of the active resistance protein only at times of pathogen attack, but would also allow a broader pathogen range to induce the hypersensitive reaction. Mutation of threonine and serine residues to aspartic acid and glutamic acid residues frequently leads to activation, as was shown in many proteins of which the activity is modulated by phosphorylation, e.g. in a MAPK-

activated protein (Engel et al., 1995, J. Biol. Chem. 270, 27213-27221), and in a MAP-kinase-kinase protein (Huang et al.,1995 Mol. Biol.Cell 6, 237-245). Also C- and N-terminal as well as internal deletion mutants of these proteins can be tested for suitable mutants.

A more undirected way of identifying interesting mutants of which constitutive activity is induced is through propagation of the protein-encoding DNA in so-called E. coli 'mutator' strains.

A rapid way of testing all made mutants for their suitability to elicit a hypersensitive response is through a so-called ATTA assay (Van den Ackerveken, G., et al., Cell 87, 1307-1316, 1996). Many mutants can be screened with low effort to identify those that will elicit an HR upon expression.

The invention also provides a vector comprising a nucleic acid as provided herein, said nucleic acid encoding a gene product that is capable of providing a member of the *Solanaceae* family with resistance against an oomycete pathogen, or a functionally equivalent isolated or recombinant nucleic acid in particular wherein said member comprises *S. tuberosum* or *Lycopersicon esculentum*.

The invention also provides a host cell comprising a nucleic acid or a vector according to the invention. An example of said host cell is provided in the detailed description herein. In a particular embodiment, said host cell comprises a plant cell. As a plant cell a cell derived from a member of the *Solanaceae* family is preferred, in particular wherein said member comprises *S. tuberosum* or *Lycopersicon esculentum*. From such a cell, or protoplast, a transgenic plant, such as transgenic potato plant or tomato plant with resistance against an oomycete infection can arise. The invention thus also provides a plant, or tuber root, fruit or seed or part or progeny derived thereof comprising a cell according to the invention.

Furthermore, the invention provides a proteinaceous substance, exhibiting the characteristic of providing at least partial resistance to an oomycete infection such as caused by *P. infestans* when incorporated and expressed in a plant or plant cell. In particular such a proteinaceous substance is provided that is encoded by a nucleic acid according to the invention. In a preferred embodiment, the invention provides a proteinaceous substance comprising an amino acid sequence as depicted in figure 8 or a functional equivalent thereof. Preferably, such a functional equivalent will comprise one or more sequences which are relatively unique to *Rp1-blb* in comparison to

RGC3-blb, RGC-blb and *RGC4-blb*. Such sequences can be spotted in the alignment (see fig. 10A) and would be the sequences RPLLGEM, AKMEKEKLIS, KHSYTHMM, FFYTLPPLEKFI, GDSTFNK, NLYGSGMRS, LQYCTKLC, GSQSLTCM, NNFGPHI, TSLKIYGFRGIH, IIHECPFLTLS, RICYNKVA, and KYLTISRCN. It is believed that one or more of these sequences provide the functional characteristics of the protein *Rp1-blb*.

Furthermore, the invention provides a binding molecule directed at a nucleic acid according to the invention. For example, the *Rpi-blb* gene can be used for the design of oligonucleotides complementary to one strand of the DNA sequence as depicted in Figure 7 and Table 2. Such oligonucleotides as provided herein are useful as probes for library screening, hybridisation probes for Southern/Northern analysis, primers for PCR, for use in a diagnostic kit for the detection of disease resistance and so on. Such oligonucleotides are useful fragments of an isolated or recombinant nucleic acid as provided herein, said nucleic acid encoding a gene product that is capable of providing a member of the *Solanaceae* family with resistance against an oomycete fungus, or a functionally equivalent isolated or recombinant nucleic acid, in particular wherein said member comprises *S. tuberosum* or *Lycopersicon esculentum*. They can be easily selected from a sequence as depicted in figure 6 or part thereof. A particular point of recognition comprises the LRR domain as identified herein. Such a binding molecule according to the invention is used as a probe or primer, for example provided with a label, in particular wherein said label comprises an excitable moiety which makes it useful to detect the presence of said binding molecule.

The invention furthermore provides a method for selecting a plant or plant material or progeny thereof for its susceptibility or resistance to an oomycete infection comprising testing at least part of said plant or plant material or progeny thereof for the presence or absence of a nucleic acid, said nucleic acid encoding a gene product that is capable of providing a member of the *Solanaceae* family with resistance against an oomycete fungus, or for the presence of said gene product, said method preferably comprising contacting at least part of said plant or plant material or progeny thereof with a binding molecule according the invention and determining the binding of said molecule to said part. Said method is particularly useful wherein said oomycete comprises *P. infestans*, allowing to select plants or planting material for resistance against late blight, for example wherein said plant or material comprises *S. tuberosum*. It is believed that by the phylogenetic tree analysis as discussed above, proteins that are highly homologous to *Rpi-blb* and which would

yield resistance against plant pathogens could be easily idientified. An example for this is the detection of the three highly homologous proteins *RGC1-blb, RGC3-blb* and *RGC4-blb*, which have not yet been shown to yield resistance to *P. infestans*, but which are nevertheless believed to be involved in pathogen resistance in plants.

Also, the invention provides use of a nucleic acid or a vector or a cell or a substance or a binding molecule according to the invention in a method for providing a plant or its progeny with at least partial resistance against an oomycete infection, in particular wherein said oomycete comprises *P. infestans* especially wherein said plant comprises *S. tuberosum*, said method for providing a plant or its progeny with at least partial resistance against an oomycete infection comprising providing said plant or part thereof with a gene coding for a resistance protein or functional fragment thereof comprising a nucleic acid, said resistance protein being capable of providing a member of the *Solanaceae* family with resistance against an oomycete fungus, or providing said plant or part thereof with a nucleic acid or a vector or a cell or a substance according to the invention.

Furthermore, the invention provides an isolated *S. bulbocastanum*, or part thereof, such as a tuber or seed, susceptible to an oomycete infection caused by *P. infestans*.

The invention is further described in the detailed description below.

DESCRIPTION OF THE FIGURES

Figure 1. Geographical map of Mexico indicating the origin of *Solanum bulbocastanum* accessions used to isolate the *Rpi-blb* gene. The letters a, b and c indicate the relative geographical origins of the used *S. bulbocastanum* accessions.

Figure 2. Genetic linkage maps of the *Rpi-blb* locus on chromosome 8 of *S. bulbocastanum*. Horizontal lines indicate the relative positions of markers linked to late blight resistance. Distances between markers are indicated in centimorgans. **A.** Genetic position of the *Rpi-blb* locus relative to markers TG513, CT88 and CT64 (n=508 genotypes). **B.** High density genetic linkage map of the *Rpi-blb* locus (n=2109 genotypes).

Figure 3. Physical map of the *Rpi-blb* locus. **A.** Genetic and physical map of the *S. bulbocastanum* genomic region containing *Rpi-blb*. Vertical arrows indicate the relative positions of markers linked to resistance. Numbers above the horizontal line indicate the number of recombinants identified between the flanking markers in 2109 progeny plants. Rectangles represent bacterial artificial chromosome (BAC) clones. **B.** Relative positions of candidate genes for late blight resistance on BAC SPB4. **C.** Schematic representation of the *Rpi-blb* gene structure. Horizontal lines indicate exons. Open boxes represent coding sequence. Lines angled downwards indicate the position of a 678-nucleotide long intron sequence.

Figure 4. Southern blot analysis of the BAC contig spanning the *Rpi-blb* locus. Names above each lane represent the names of BAC clones. The names of the restriction enzymes used to digest the BAC DNA prior to Southern blotting are indicated.

Figure 5. Detached leaf disease assays. **A.** Resistant (left), intermediate (centre) and susceptible (right) phenotypes found in the *S. bulbocastanum* mapping population B8 6 days post inoculation (d.p.i) with *P. infestans* sporangiospore droplets. **B.** Genetic complementation for late blight resistance in potato. Characteristic disease phenotypes of leaves derived from transgenic potato plants harbouring *RGC1-blb*, *RGC2-blb*, *-blb* or *RGC4-blb* 6 d.p.i. with *P. infestans* sporangiospore droplets. Genetic constructs harbouring the RGCs were transferred to the susceptible potato cultivar Impala through *Agrobacterium* mediated transformation. **C.** Genetic

complementation for late blight resistance in tomato. Characteristic disease phenotype of a tomato leaf derived from transgenic tomat plants harbouring *Rpi-blb* 6 d.p.i. with *P. infestans* sporangiospore droplets (left panel). The genetic construct harbouring *Rpi-blb* was transferred to the susceptible tomato cultivar Moneymaker through *Agrobacterium* mediated transformation.

Figure 6. Nucleic acid sequences of the *Rpi-blb* gene cluster members. **A.** Coding nucleic acid sequence of the *Rpi-blb* gene. **B.** Coding nucleic acid sequence of the *Rpi-blb* gene including the intron sequence (position 428-1106). **C.** Sequence of the 5.2 kb *Sca*I genomic DNA fragment of *S. bulbocastanum* BAC SPB4 present in pRGC2-blb, the genetic construct used for genetic complementation for late blight resistance. The genomic fragment harbours the *Rpi-blb* gene including natural regulatory elements necessary for correct expression of the gene. The initiation codon (ATG position 1191-1193) and the termination codon (TAA position4781-4783) are underlined. **D.** Coding nucleic acid sequence of *RGC1-blb* including the intron sequence (position 428-708). **E.** Coding nucleic acid sequence of *RGC3-blb* including the intron sequence (position 428-1458). **F.** Coding nucleic acid sequence of *RGC4-blb* including intron sequences (positions 434-510, 543-618 and 743-1365).

Figure 7. Relative primer positions. The horizontal bar represents the coding sequence of the *Rpi-blb* gene. Numbers represent nucleotide positions. Horizontal arrows indicate relative primer positions and orientations. GSP1 and GSP2 represent nested gene specific primers used for 3' RACE experiments. GSP3 and GSP4 represent nested gene specific primers used for 5' RACE experiments. A(F), A(R), B(F) and B(R) are primers used to amplify *Rpi-blb* homologues. The position of the restriction site *Nsi*I used to make domain swaps between *Rpi-blb* homologues is indicated.

Figure 8. Deduced Rpi-blb protein sequence. The amino acid sequence deduced from the DNA sequence of *Rpi-blb* is divided into three domains (A-C), as described in Example 6. Hydrophobic residues in domain A that form the first and fourth residues of heptad repeats of potential coiled-coil domains are underlined. Conserved motifs in R proteins are written in lowercase and in italic in domain B. Residues matching the consensus of the cytoplasmic LRR are indicated in bold in domain C. Dots in the

sequence have been introduced to align the sequence to the consensus LRR sequence of cytoplasmic LRRs.

Figure 9. Phylogenetic tree analysis. **A.** Phylogenetic tree of state of the art sequences which share some degree of homology to the deduced amino acid sequence of *Rpi-blb* and its gene cluster members *RGC1-blb*, *RGC3-blb* and *RGC4-blb*. The tree was made according to the Neighbour-Joining method of Saitou and Nei (1987 Molecular Biology and Evolution 4, 406-425). An asterix indicates that the gene has been assigned a function. The *Rpi-blb* gene cluster is boxed. **B.** Phylogenetic tree of state of the art sequences which share some degree of homology to the deduced amino acid sequence of *Rpi-blb*. Included in this analysis are the *Rpi-blb* homologous sequences *B149-blb*, *SH10-tub*, *SH20-tub* and *T118-tar*, sequences identified through PCR amplification using *Rpi-blb* gene cluster specific primers. **C.** Relative positions of state of the art DNA sequences which show significant homology to parts of the *Rpi-blb* gene sequence. Horizontal lines represent the relative positions of the homologous sequences. The degree of homology is indicated to the right of each line. The length of the homologous sequence is indicated above each line.

Figure 10. Alignment of the predicted *Rpi-blb* gene product to the predicted protein sequences of *Rpi-blb* homologues **A.** Alignment of the deduced protein products encoded by *Rpi-blb*, *RGC1-blb*, *RGC3-blb* and *RGC4-blb*. The complete amino acid sequence of Rpi-blb is shown and amino acid residues from RGC1-blb, RGC3-blb and RGC4-blb that differ from the corresponding residue in Rpi-blb. Dashes indicate gaps inserted to maintain optimal alignment. Amino acid residues that are specific for Rpi-blb, when compared to those at corresponding positions in RGC1-blb, RGC3-blb and RGC4-blb, are highlighted in bold. The regions of the LRRs that correspond to the consensus L..L..L.L..C/N/S..a..aP are underlined. Conserved motifs in the NBS domain are indicated in lowercase. **B.** Alignment of the deduced protein products encoded by *Rpi-blb*, *RGC1-blb*, *RGC3-blb*, *RGC4-blb*, *B149-blb*, *SH10-tub*, *SH20-tub* and *T118-tar*..

Figure 11. Schematic overview of domain swaps made between *Rpi-blb* and homologues *RGC1-blb* and *RGC3-blb*. The vertical dotted line indicates the position of the *NsiI* site used to make the swaps. R and S indicate whether transgenic plants harbouring specific chimeric constructs are resistant or susceptible to late blight infection, respectively.

Experimental part

For the mapping of the *Rpi-blb* resistance gene an intraspecific mapping population of *S. bulbocastanum* was developed. A crucial step in this process was the identification of susceptible *S. bulbocastanum* genotypes. For this purpose several *S. bulbocastanum* accessions originating from different clusters/areas in Mexico were analysed for *P. infestans* resistance or susceptibility in a detached leaf assay (Table 1 and Figure 1). The screened accessions BGRC 8008 and BGRC 7999 contained no susceptible genotypes. However in the accessions BGRC 8005, BGRC 8006 and BGRC 7997, susceptibility was found in 9%, 7% and 14 % of the analysed seedlings, respectively. A *P. infestans* susceptible clone of accession BGRC 8006 was subsequently selected and crossed with a resistant clone of accession BGRC 8005. The resulting F1 population was used to map the *Rpi-blb* locus and is hereafter referred to as the B8 population.

Initial screening of 42 B8 genotypes for resistance to *P. infestans* in a detached leaf assay suggested that *P. infestans* resistance in *S. bulbocastanum* accession 8005 could be caused by a single dominant *R* gene, or a tightly linked gene cluster. Of the 42 genotypes tested, 22 scored resistant and 16 susceptible in a repeated experiment. Resistance phenotypes of the remaining 4 seedlings remained unclear. In order to determine the chromosome position of this *S. bulbocastanum* resistance, B8 genotypes with an undoubted phenotype were used for marker analysis. The chromosome 8 specific marker TG330 (Table 2) was found to be linked in repulsion phase with the resistant phenotype, as only one recombinant was obtained between this marker and resistance in 12 B8 genotypes. Furthermore, chromosome 8 marker CT88 (Table 2) was found to be completely linked in repulsion phase to resistance, indicating that the locus responsible for resistance, designated *Rpi-blb*, was located in this region of chromosome 8. For this reason, tomato chromosome 8 specific markers that map proximal and distal to CT88 (TG513 and CT64; Tanksley et al., 1992 Genetics 132: 1141-1160; Table 2) were developed into CAPS markers and tested in 512 B8 genotypes with known resistance phenotypes. A total of five CT64-CT88 recombinant genotypes and 41 CT88-TG513 recombinant genotypes were identified in this screen (Figure 2A). The resistance locus *Rpi-blb* was mapped 1 recombination event distal to marker CT88 (Figure 2A).

Fine mapping of the *Rpi-blb* locus was carried out with CAPS markers derived from left (L) and right (R) border sequences of BAC clones isolated from a BAC library prepared from the resistant *S. bulbocastanum* genotype BGRC 8005-8. The BAC library was initially screened with markers CT88 and CT64. BAC clones identified with these markers were used as seed BACs for a subsequent chromosome walk to the *Rpi-blb* locus. A total of 2109 B8 genotypes were screened for recombination between markers TG513 en CT64. All recombinant genotypes (219/2109) were subsequently screened with all available markers in the CT88-CT64 genetic interval. These data together with the disease resistance data of each recombinant, obtained through detached leaf assays, positioned the *Rpi-blb* locus between markers SPB33L and B149R, a 0.1 cM genetic interval (4/2109 recombinants) physically spanned by the overlapping BAC clones SPB4 and B49 (Figures 2b and 3). Within this interval resistance cosegregated with the BAC end marker SPB42L, the sequence of which was highly homologous to partial NBS fragments from tomato (e.g. Q194, Q137, Q152, Q153;Pan et al., 2000 Genetics 155: 309-322). Southern analyses of BAC clones spanning the SP33L-B149R interval using a ^{32}P-labeled PCR fragment of marker SPB42L as a probe revealed the presence of at least 4 copies of this *R* gene like sequence within the *Rpi-blb* interval (Figure 4). Moreover, all of these copies were present on BAC SPB4. Sequencing and annotation of the complete insert of this BAC clone indeed identified four complete *R* gene candidates (*RGC1-blb, RGC2-blb, RGC3-blb and RGC4-blb*) of the NBS-LRR class of plant *R* genes. A PCR-marker that was located in-between *RGC1-blb* and *RGC4-blb* revealed recombination between *P. infestans* resistance and *RGC4-blb*, ruling out the possibility of *RGC4-blb* being *Rpi-blb*. Despite this finding, all four RGCs were selected for complementation analysis.

Genomic fragments of approximately 10 kb harbouring *RGC1-blb*, *RGC2-blb*, *RGC3-blb* or *RGC4-blb* were subcloned from BAC SPB4 into the binary plant transformation vector pBINPLUS (van Engelen et al., 1995 Trans. Res. 4, 288-290) and transferred to a susceptible potato cultivar using standard transformation methods. Primary transformants were tested for *P. infestans* resistance as described in Example 1. Only the genetic construct harbouring *RGC2-blb* was able to complement the susceptible phenotype; 86% of the primary transformants harbouring *RGC2-blb* were resistant (Table 3) whereas all *RGC1-blb*, *RGC3-blb* and *RGC4-blb* containing primary transformants were completely susceptible to *P. infestans*. The resistant *RGC2-blb* containing transformants showed similar resistance phenotypes

as the *S. bulbocastanum* resistant parent (Figure 5). *RGC2-blb* was therefore designated the *Rpi-blb* gene, the DNA sequence of which is provided in Figure 6.

EXAMPLE 1: DISEASE ASSAY

The phenotype of *S. bulbocastanum* and transgenic *S. tuberosum* genotypes for resistance to *P. infestans* was determined by detached leaf assays. Leaves from plants grown for 6 to 12 weeks in the greenhouse were placed in pieces of water-saturated florists foam, approximately 35x4x4 cm, and put in a tray (40 cm width, 60 cm length and 6 cm height) with a perforated bottom. Each leaf was inoculated with two droplets or more (25 µl each) of sporangiospore solution on the abaxial side. Subsequently, the tray was placed in a plastic bag on top of a tray, in which a water-saturated filter paper was placed, and incubated in a climate room at 17°C and a 16h/8h day/night photoperiod with fluorescent light (Philips TLD50W/84HF). After 6 days, the leaves were evaluated for the development of *P. infestans* disease symptoms. Plants with leaves that clearly showed sporulating lesions 6 days after inoculation were considered to have a susceptible phenotype whereas plants with leaves showing no visible symptoms or necrosis at the side of inoculation in the absence of clear sporulation were considered to be resistant. The assay was performed with *P. infestans* complex isolate 655-2A (race 1, 2, 3,4, 5, 6, 7, 8, 9, 10, 11), which was obtained from Plant Research International BV (Wageningen, The Netherlands).

EXAMPLE 2: MAPPING OF THE *Rpi-blb* RESISTANCE LOCUS

Plant material

In order to produce an intraspecific mapping population that segregated for the *P. infestans* resistance gene present in *S. bulbocastanum* accession BGRC 8005 (CGN 17692, PI 275193), a susceptible *S. bulbocastanum* genotype was required. Several *S. bulbocastanum* accessions originating from different clusters/areas in Mexico were analysed for *P. infestans* resistance or susceptibility in a detached leaf assay (Table 1 and Figure 1). In accession BGRC 8008 and BGRC 7999 no susceptibility was detected. In accession BGRC 8005, BGRC 8006 and BGRC 7997 susceptibility was only present in 9%, 7% and 14 % of the analysed seedlings, respectively. Thus, only a few susceptible *S. bulbocastanum* genotypes were obtained.

The intraspecific mapping population of *S. bulbocastanum* (B8) was produced by crossing a *P. infestans* susceptible clone of accession BGRC 8006 with a resistant clone of accession BGRC 8005. DNA of 2109 progeny plants was extracted from young leaves according to Doyle and Doyle (1989 Focus 12, 13-15).

CAPS marker analysis

For PCR analysis, 15 µl reaction mixtures were prepared containing 0.5 µg DNA, 15 ng of each primer, 0.2 mM of each dNTP, 0.6 units Taq-polymerase (15 U/µl, SphaeroQ, Leiden, The Netherlands), 10 mM Tris-HCl pH 9, 1.5 mM $MgCl_2$, 50 mM KCl, 0.1% Triton X-100 and 0.01% (w/v) gelatin. The PCRs were performed using the following cycle profile: 25 seconds DNA denaturation at 94°C, 30 seconds annealing (see Table 1) and 40 seconds elongation at 72°C. As a first step in PCR-amplification DNA was denatured for 5 min at 94°C and finalised by an extra 5 min elongation step at 72°C. The amplification reactions were performed in a Biometra® T-Gradient or Biometra® Uno-II thermocycler (Westburg, Leusden, The Netherlands). Depending on the marker, the PCR product was digested with an appropriate restriction enzyme. An overview of the markers including primer sequences, annealing temperature and restriction enzymes, is given in Table 2. Subsequently, the (digested) PCR products were analysed by electrophoresis in agarose or acrylamide gels. For acrylamide gel analysis, the CleanGel DNA Analysis Kit and DNA Silver Staining Kit (Amersham Pharmacia Biotech Benelux, Roosendaal, the Netherlands) were used.

Genetic mapping of the *Rpi-blb* locus

Initially a small group of 42 progeny plants of the B8 population was screened for resistance to *P. infestans* in a detached leaf assay. Plants with leaves that clearly showed sporulating lesions 6 days after inoculation were considered to have a susceptible phenotype whereas plants with leaves showing no visible symptoms or necrosis at the side of inoculation in the absence of clear sporulation were considered to be resistant. Of the 42 seedlings, 22 scored resistant and 16 susceptible. The phenotype of the remaining 4 seedlings remained unclear in this initial phase. These data indicated that resistance could be due to a single dominant gene or a tightly linked gene cluster. In order to determine the chromosome position, seedlings with a reliable phenotype were used for marker analysis. Chromosome 8 marker TG330 was found to be linked in repulsion with the resistant phenotype, as only one recombinant was obtained between this marker and resistance in 12 B8 seedlings. Furthermore,

chromosome 8 marker CT88 was found to be completely linked in repulsion phase to resistance, indicating that a resistance gene was located on chromosome 8.

Subsequently, chromosome 8 specific markers that had been mapped proximal and distal to CT88 (Tanksley et al., 1992 Genetics 132: 1141-1160) were developed to CAPS markers. In order to map these markers more precisely, another 512 individuals of the B8 population were screened for late blight resistance using the detached leaf disease assay. Simultaneously, plants were scored for the markers CT64, CT88 and TG513. For 5 seedlings, recombination was detected between markers CT64 and CT88, while 41 seedlings were recombinant between markers CT88 and TG513 (Figure 2A). The resistance gene *Rpi-blb* was mapped in between markers CT64 and CT88. In this stage, the positioning of CT88 proximal to *Rpi-blb* was based on only one recombined seedling.

In order to determine the position of *Rpi-blb* more precisely relative to the available markers, another 1555 seedlings of the B8 population were grown and analysed for recombination between the markers TG513 and CT64. Thus, a total of 2109 individual offspring clones of the B8 population were screened. Recombination between markers TG513 en CT64 was detected in 219 of these seedlings (10.4 cM). All of the recombinants were screened with marker CT88 and phenotyped for the resistance trait by making use of the detached leaf assay. In agreement with earlier results, the *Rpi-blb* gene was mapped in between markers CT88 and CT64 (Figure 2B).

EXAMPLE 3: CONSTRUCTION OF A *S. BULBOCASTANUM* BAC LIBRARY AND CONSTRUCTION OF A CONTIGUOUS BAC CONTIG SPANNING THE *Rpi-blb* LOCUS

BAC library construction

A resistant clone of *S. bulbocastanum* (blb) accession BGRC 8005 (CGN 17692, PI 275193) heterozygous for the *Rpi-blb* locus, was used as source DNA for the construction of a genomic BAC library, hereafter referred to as the 8005-8 BAC library. High molecular weight DNA preparation and BAC library construction were carried out as described in Rouppe van der Voort et al. (1999 MPMI 12:197-206). Approximately 130.000 clones with an average insert size of 100 kb, which corresponds to 15 genome equivalents were finally obtained. A total of approximately

83.000 individual clones were stored in 216 384-well microtiter plates (Invitrogen, The Netherlands) containing LB freezing buffer (36 mM K_2HPO_4, 13.2 mM KH_2PO_4, 1.7 mM citrate, 0.4 mM $MgSO_4$, 6.8 mM $(NH_4)_2SO_4$, 4.4 % V/V glycerol, 12.5 µg/ml chloramphenicol in LB medium) at -80°C. Another 50.000 clones were stored as bacterial pools containing ~1000 white colonies. These were generated by scraping the colonies from the agar plates into LB medium containing 18% glycerol and 12.5 µg/ml chloramphenicol using a sterile glass spreader. These so-called super pools were also stored at -80°C.

Screening of the BAC library and construction of a physical map of the *Rpi-blb* locus

The 8005-8 BAC library was initially screened with CAPS markers CT88 and CT64. This was carried out as follows. For the first part of the library of approximately 83.000 clones stored in 384 well microtiter plates, plasmid DNA was isolated using the standard alkaline lysis protocol (Sambrook *et al.*, 1989 in Molecular cloning: a laboratory manual 2nd edn, Cold Spring Harbor Press, Cold Spring Harbor, New York) from pooled bacteria of each plate to produce 216 plate pools. To identify individual BAC clones carrying the CAPS markers the plate pools were screened by PCR. Once an individual plate pool was identified as being positive for a particular CAPS marker the positive row and positive column were identified through a two dimensional PCR screening. For this purpose, the mother 384-well plate was replicated twice on LB medium containing chloramphenicol (12.5 µg/ml). After growing the colonies for 16 h at 37°C one plate was used to scrape the 24 colonies of each row together and the other plate was used to scrape the 16 colonies of each column together. Bacteria of each row or column were resuspended in 200 µl TE buffer. CAPS marker analysis on 5 µl of these bacterial suspensions was subsequently carried out leading to the identification of single positive BAC clones. For the second part of the library, stored as 50 pools of approximately 1000 clones, plasmid DNA was isolated from each pool of clones using the standard alkaline lysis protocol and PCR was carried out to identify positive pools. Bacteria corresponding to positive pools were diluted and plated on LB agar plates containing chloramphenicol (12.5 µg/ml). Individual white colonies were subsequently picked into 384-well microtiter plates and single positive BAC clones subsequently identified as described above. Names of BAC clones isolated from the super pools carry the prefix SP (e.g. SPB33).

Insert sizes of BAC clones were estimated as follows. Positive BAC clones were analysed by isolating plasmid DNA from 2 ml overnight cultures (LB medium supplemented with 12.5 mg/ml chloramphenicol) using the standard alkaline lysis miniprep protocol and resuspended in 20 µl TE. Plasmid DNA (10 µl) was digested with 5 U *Not*I for 3 h at 37°C to free the genomic DNA from the pBeloBAC11 vector. The digested DNA was separated by CHEF electrophoresis in a 1% agarose gel in 0.5 X TBE at 4°C using a BIORAD CHEF DR II system (Bio-Rad Laboratories, USA) at 150 volts with a constant pulse time of 14 sec for 16 h.

Screening of the 8005-8 BAC library with marker CT88 identified two positive BAC clones: B139 and B180, with potato DNA inserts of 130 and 120 kb, respectively (Figure 3A). Digestion of the CT88 PCR product generated from these BAC clones and several resistant and susceptible progeny plants of the B8 mapping population with *Mbo*I revealed that BAC139 carried the CT88 allele that was linked in *cis* to resistance. To identify the relative genome position of BAC B139, pairs of PCR primers were designed based on the sequence of the right (R) and left (L) ends of the insert. BAC end sequencing was carried out as described in Example 4 using 0.5 µg of BAC DNA as template. Polymorphic CAPS markers were developed by digesting the PCR products of the two parent genotypes of the B8 population and of two resistant and two susceptible progeny genotypes with several 4-base cutting restriction enzymes (Table 2). Screening of the 37 CT88-CT64 recombinant B8 genotypes mapped 5 of the 7 CT88-*Rpi-blb* recombinants between CT88 and B139R, indicating that marker B139R was relatively closer to the *Rpi-blb* locus than marker CT88. Screening of the 216 plate pools with B139R did not lead to the identification of a positive BAC clone. Screening of the 50 super pools identified the positive BAC clones SPB33 and SPB42 with DNA inserts of 85 and 75 kb, respectively (Figure 3A). Screening of the complete BAC library with SPB33L identified the positive BAC clones B149 and SPB4. BAC clone SPB4 contained the SPB33L allele that was linked *in cis* to resistance whereas BAC clone B149 did not. However, screening of the CT88-CT64 recombinant panel with B149R revealed that this BAC spanned the *Rpi-blb* locus. B149R was separated from the *Rpi-blb* locus by two recombination events (Figure 3A). Screening of the 8005-8 BAC library with B149R identified BAC clone B49 as having the B149R allele that was linked *in cis* to resistance. This BAC clone together with BAC clone SPB4 therefore formed a BAC contig that spanned the *Rpi-blb* locus (Figure 3).

EXAMPLE 4: SEQUENCE ANALYSIS OF BAC SPB4 AND IDENTIFICATION OF RESISTANCE GENE CANDIDATES WITHIN THE *Rpi-blb* LOCUS

Within the SPB33L-B149R interval resistance cosegregated with BAC end marker SPB42L, the sequence of which was highly homologous to partial NBS fragments from tomato (e.g. Q194, Q137, Q97, Q152, Q153; Pan et al., 2000 Genetics 155:309-22). Southern analyses of BAC clones spanning the SPB33L-B149R interval using a ^{32}P-labeled PCR fragment of marker SPB42L as a probe revealed the presence of at least 4 copies of this *R* gene like sequence within the *Rpi-blb* interval (Figure 4). Moreover, all of these copies were present on BAC SPB4. The DNA sequence of BAC clone SPB4 was therefore determined by shotgun sequence analysis. A set of random subclones with an average insert size of 1.5 kb was generated. 10 µg of CsCl purified DNA was sheared for 6 seconds on ice at 6 amplitude microns in 200 µl TE using an MSE soniprep 150 sonicator. After ethanol precipitation and resuspension in 20 µl TE the ends of the DNA fragments were repaired by T4 DNA polymerase incubation at 11°C for 25 minutes in a 50 µl reaction mixture comprising 1x T4 DNA polymerase buffer (New England BioLabs, USA), 1 mM DTT, 100 µM of all 4 dNTP's and 25 U T4 DNA polymerase (New England Biolabs, USA), followed by incubation at 65°C for 15 minutes. The sheared DNA was subsequently separated by electrophoresis on 1% SeaPlaque LMP agarose gel (FMC). The fraction with a size of 1.5-2.5 kb was excised from the gel and dialysed against 50 ml TE for 2 hr at 4°C. Dialysed agarose slices were then transferred to a 1.5 ml Eppendorf tube, melted at 70°C for 5 min, digested with 1 unit of GELASE (Epicentre Technologies, USA) per 100 mg of agarose gel for 1 hr at 45°C, and the DNA was subsequently precipitated. The 1.5-2.5 kb fragments were ligated at 16°C in a *Eco*RV restricted and dephosphorylated pBluescript SK$^+$ vector (Stratagene Inc.). The ligation mixture was subsequently used to transform ElectroMAX *E. coli* DH10B competent cells (Life Technologies, UK) by electroporation using the BioRad Gene Pulser. Settings on the BioRad Gene Pulser were as recommended for *E. coli* by the manufacturer. The cells were spread on Luria broth (LB) agar plates containing ampicillin (100 µg/ml), 5-bromo-4-chloro-3-indolyl-ß-D-galactoside (Xgal) (64 µg/ml) and isopropyl-1-thio-ß-D-galactoside (IPTG) (32 µg/ml). Plates were incubated at 37°C for 24 hours. Individual white colonies were grown in 96-well flat-bottom blocks (1.5 ml Terrific Broth medium containing 100 µg/ml ampicillin).

Plasmid DNA was isolated using the QIAprep 96 Turbo Miniprep system in conjunction with the BioRobot™ 9600 (QIAGEN) according to the manufacturers instructions. Sequencing reactions were performed using ABI PRISM BigDye™ Terminator cycle sequencing kit (Stratagene) according to the manufacturer's instructions. All clones were sequenced bi-directionally using universal primers. Sequence products were separated by capillary electrophoresis on a Perkin Elmer ABI 3700 DNA Analyzer.

The automated assembly of the shotgun reads was carried out using the Phred-Phrap programs (Ewing and Green, 1998 Genome Research 8, 186-194; Ewing et al., 1998 Genome Research 8, 175-185). A total of 835 reads provided an overall BAC sequence coverage equal to 5x. Gaps between contigs were closed by primer walking or through a combinatorial PCR approach. The sequence was finally edited at Phred quality 40 (1 error every 10,000 nt) by manual inspection of the assembly using the Gap4 contig editor and re-sequencing of all low-quality regions. The complete sequence of the insert of BAC SPB4 consisted of 77,283 nucleotides.

Analysis of the contiguous sequence of BAC SPB4 using the computer programme GENSCAN (Burge and Karlin, 1997 J. Mol. Biol. 268, 78-94), GENEMARK (Lukashin and Borodovsky, 1998 NAR 26, 1107-1115) and BLASTX (Altschul et al., 1990 J. Mol. Biol. 215, 403-410) identified four complete R gene candidate sequences (*RGC1-blb, RGC2-blb, RGC3-blb and RGC4-blb*) belonging to the NBS-LRR class of plant *R* genes. A CAPS marker designed in between *RGC1-blb* and *RGC4-blb*, marker RGC1-4 revealed recombination between *P. infestans* resistance and *RGC4-blb*, ruling out the possibility of *RGC4-blb* being *Rpi-blb* (Figure 3A and B). Despite this finding, all four RGCs were selected for complementation analysis.

EXAMPLE 5: COMPLEMENTATION ANALYSIS

Subcloning of candidate genes and transformation to *Agrobacterium tumefaciens*

Genomic fragments of approximately 10 kb harbouring *RGC1-blb, RGC2-blb, RGC3-blb or RGC4-blb* were subcloned from BAC clone SPB4 into the binary plant transformation vector pBINPLUS (van Engelen et al., 1995 Trans. Res. 4, 288-290). Restriction enzyme digestion of BAC clone SPB4 DNA and subsequent size selection was carried out as follows. Aliquots of ~1 µg DNA were digested with 1U, 0.1U or

0.01U of *Sau3*AI restriction enzym for 30 min. The partially digested BAC DNA was subjected to CHEF electrophoresis at 4°C in 0.5 X TBE using a linear increasing pulse time of 1-10 sec and a field strength of 6 V/cm for 16 hr. After electrophoresis, the agarose gel was stained with ethidium bromide to locate the region of the gel containing DNA fragments of approximately 10kb in size. This region was excised from the gel using a glass coverslip and dialysed against 50 ml TE for 2 hr at 4°C. Dialysed agarose slices were then transferred to a 1.5 ml Eppendorf tube, melted at 70°C for 5 min and digested with 1 unit of GELASE (Epicentre Technologies, USA) per 100 mg of agarose gel for 1 hr at 45°C. Ligation of the size selected DNA to *Bam*HI-digested and dephosphorylated pBINPLUS and subsequent transformation of ElectroMAX *E. coli* DH10B competent cells (Life Technologies, UK) with the ligated DNA was carried as described in Example 5, using the BioRad Gene Pulser for electroporation. The cells were spread on Luria broth (LB) agar plates containing kanamycin (50 µg/ml), Xgal (64 µg/ml) and IPTG (32 µg/ml). Plates were incubated at 37°C for 24 hours. Individual white colonies were grown in 96-well plates (100 µl LB medium containing 50 µg/ml kanamycin). A total of 480 clones were PCR screened for the presence of RGCs using primers SPB42LF and SPB42LR or RGC4F and RGC4R (Table 2.). Positive clones were selected for plasmid isolation and further characterisation. Identification of clones harbouring *RGC1-blb, RGC2-blb, RGC3-blb* or *RGC4-blb* was carried out by sequencing the SPB42L PCR fragments derived from positive clones. The relative position of the RGCs within a subclone was determined by sequencing the ends of the clone and subsequent comparison of the sequences to the complete BAC insert sequence. Finally four binary plasmids, pRGC1-blb, pRGC2-blb, pRGC3-blb and pRGC4-blb were selected and transferred to Agrobacterium tumefaciens strains AGL0 (Lazo et al., 1991 Bio/Technology 9, 963-967), LBA4404 (Hoekema et al., 1983 Nature 303: 179-180) or UIA143 (Farrand et al.,1989 J. of Bacteriology 171, 5314-5321) either by electroporation using the BioRad Gene Pulser or by conjugation. Settings on the BioRad Gene Pulser were as recommended for *A. tumefaciens* by the manufacturer. Conjugation was carried out as described by Simon et al. (1983 Bio/Tech. 1, 784-791). The cells were spread on Luria broth (LB) agar plates containing kanamycin (100 mg/l) and rifampicin (50 mg/l). Plates were incubated at 28°C for 48 hours. Small-scale cultures from selected colonies were grown in LB medium containing kanamycin (100 mg/l) and rifampicin (50 mg/l). Plasmid DNA was isolated as described previously and the integrity of the plasmids was verified by restriction analysis upon reisolation from *A. tumefaciens* and

subsequent transformation to *E. coli*. *A tumefaciens* cultures harbouring a plasmid with the correct DNA pattern were used to transform a susceptible potato genotype.

Transformation of susceptible potato cultivar

A. tumefaciens strains were grown for 2 days at 28°C in 20 ml LB medium supplemented with 50 mg/l rifampicine and 25 mg/l kanamycin. Subsequently, 0.2 ml of *A. tumefaciens* culture was diluted in 10 ml LB medium containing the same antibiotics and grown overnight (28°C). The overnight culture was centrifuged (30 min, 2647 x g) and the pellet was resuspended in 50 ml MS medium (Murashige and Skoog, 1962 Physiol. Plant. 15, 473-497) supplemented with 30 g/l sucrose (MS30).

Certified seed potatoes of cultivar Impala were peeled and surface sterilised for 30 min. in a 1% sodium hypochlorate solution containing 0.1 % Tween-20. Tubers were then washed thoroughly in large volumes of sterile distilled water (4 times, 10 min). Discs of approximately 2 mm thickness and 7 mm in diameter, were sliced from cylinders of tuber tissue prepared with a corkborer. The tuber discs were transferred into liquid MS30 medium containing *A. tumefaciens* and incubated for 15 min. After removing the *A. tumefaciens* solution, the tuber discs were transferred to regeneration medium containing MS30, 0.9 mg/l IAA, 3.6 mg/l zeatine riboside and 8 g/l agar (Hoekema et al., 1989 Bio/Technology 7, 273-278). The plates were incubated at 24°C, 16 hour day-length (Philips TLD50W/84HF). After 48 hours of co-cultivation, the tuber discs were rinsed for 5 min in liquid MS medium including antibiotics, 200 mg/l vancomycin, 250 mg/l cefotaxim and 75 mg/l kanamycin, and transferred to regeneration medium supplemented with the same antibiotics. The plates were incubated at 24°C, 16 hour day-length (Philips TLD50W/84HF). Every three weeks, the tuber discs were transferred to fresh medium. Regenerating shoots were transferred to MS30 medium containing 75 mg/l kanamycin. Rooting shoots were propagated *in vitro* and tested for absence of *A. tumefaciens* cells by incubating a piece of stem in 3 ml LB medium (3 weeks, 37°C, 400 rpm). One plant of each transformed regenerant was transferred to the greenhouse.

Complementation of the susceptible phenotype in potato

Primary transformants were tested for P. infestans resistance as described in Example 1. Only the genetic construct harbouring *RGC2-blb* was able to complement the susceptible phenotype; 15 out of 18 *RGC2-blb* containing primary transformants were resistant (Table 3) whereas all *RGC1-blb*, *RGC3-blb* and *RGC4-blb* containing primary transformants were completely susceptible to *P. infestans*. The resistant

RGC2-blb transformants showed similar resistance phenotypes as the *S. bulbocastanum* resistant parent (Figure 5). *RGC2-blb* was therefore designated the *Rpi-blb* gene, the DNA sequence of which is provided in Figure 6.

Transformation of susceptible tomato

Seeds of the susceptible tomato line Moneymaker were rinsed in 70% ethanol to dissolve the seed coat and washed with sterile water. Subsequently, the seeds were surface-sterilised in 1.5% sodium hypochlorite for 15 minutes, rinsed three times in sterile water and placed in containers containing 140 ml MS medium pH 6.0 (Murashige and Skoog, 1962 Physiol. Plant. 15, 473-497) supplemented with 10 g/l sucrose (MS10) and 160 ml vermiculite. The seeds were left to germinate for 8 days at 25°C and 0.5 W/M^2 light.

Eight day old cotyledon explants were pre-cultured for 24 hours in Petri dishes containing a two week old feeder layer of tobacco suspension cells plated on co-cultivation medium (MS30 pH 5.8 supplemented with Nitsch vitamines (Duchefa Biochemie BV, Haarlem, The Netherlands), 0.5 g/l MES buffer and 8 g/l Daichin agar).

Overnight cultures of *A. tumefaciens* were centrifuged and the pellet was resuspended in cell suspension medium (MS30 pH 5.8 supplemented with Nitsch vitamines, 0.5 g/l MES buffer, pH 5.8) containing 200 µM acetosyringone to a final O.D.$_{600}$ of 0.25. The explants were then infected with the diluted overnight culture of *A. tumefaciens* strain UIA143 (Farrand et al., 1989 J. of Bacteriology 171, 5314-5321) containing the helper plasmid pCH32 (Hamilton et al., 1996 PNAS 93, 9975-9979) and pRGC2-blb for 25 minutes, blotted dry on sterile filter paper and co-cultured for 48 hours on the original feeder layer plates. Culture conditions were as described above.

Following the co-cultivation, the cotyledons explants were transferred to Petri dishes with selective shoot inducing medium (MS pH 5.8 supplemented with 10 g/l glucose, including Nitsch vitamines, 0.5 g/l MES buffer, 5 g/l agargel, 2 mg/l zeatine riboside, 400 mg/l carbenicilline, 100 mg/l kanamycine, 0.1 mg/l IAA) and cultured at 25°C with 3-5 W/m^2 light. The explants were sub-cultured every 3 weeks onto fresh medium. Emerging shoots were dissected from the underlying callus and transferred to containers with selective root inducing medium (MS10 pH 5.8 supplemented with Nitsch vitamines, 0.5 g/l MES buffer, 5 g/l agargel, 0.25 mg/l IBA, 200 mg/l carbenicillin and 100 mg/l kanamycine).

Complementation of the susceptible phenotype in tomato

To investigate whether *Rpi-blb* could complement the susceptible phenotype in tomato, primary transformants of Moneymaker harbouring the *Rpi-blb* gene construct were initially challenged with the potato derived *P. infestans* isolates IPO655-2A and IPO428. Seven out of nine primary transformants were resistant (Table 3). In view of the observation that the tested potato *P. infestans* isolates were less virulent on tomato than on potato, the primary transformants were also tested with a *P. infestans* isolate collected from susceptible home garden tomato plants. Even though this isolate was significantly more virulent on Moneymaker than the previously tested ones, all 7 primary transformants remained resistant. These results illustrate the potential effectiveness of the *Rpi-blb* gene not only against complex isolates derived from potato but also to those specialised on tomato.

Molecular analysis of primary transformants

RT-PCR analysis

In order to produce cDNA, a mix of 19 µl containing 1 µg of total or polyA RNA, 0.25 mM of each dNTP, 50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT and 530 ng oligo d(T) primer, GCTGTCAACGATACGCTACGTAACGGCATGACAGTG$(T)_{18}$ was denatured (1 min 83°C). Subsequently, the mix was placed at 42°C and 1 µl reverse transcriptase (M-MLV reverse transcriptase, Promega Benelux b.v., Leiden, The Netherlands) was added. After 60 min, the mix was heated for 1 min at 99°C and transferred to ice. 2 µl cDNA was used for standard PCR.

Rapid amplification of cDNA ends

The 5' and 3' ends of the *Rpi-blb* cDNA were determined by rapid amplification of cDNA ends (RACE) using the GeneRacer™ kit (Invitrogen™, The Netherlands). 3' RACE was carried out with the primers GSP1 (5'-GAGGAATCCATCTCCCAGAG) and GSP2 (5'-GTGCTTGAAGAGATGATAATTCACGAG) in combination with the GeneRacer™ 3' primer and GeneRacer™ 3' nested primer. 5' RACE was carried out on cDNA synthesised with the primer GSP3 (5'-GTCCATCTCACCAAGTAGTGG) using primers GSP4 (5'-GAAATGCTCAGTAACTCTCTGG) and GSP5 (5'-

GGAGGACTGAAAGGTGTTGG) in combination with the GeneRacer™ 5' primer and GeneRacer™ 5' nested primer (Figure 7).

EXAMPLE 6: STRUCTURE OF THE *Rpi-blb* GENE AND THE CORRESPONDING PROTEIN.

The size and structure of the *Rpi-blb* gene was determined by comparing the genomic sequence derived from the insert of pRGC2-blb with cDNA fragments generated by 5' and 3' rapid amplification of cDNA ends. RACE identified 5' and 3' *Rpi-blb* specific cDNA fragments of a single species, respectively, suggesting that the genomic clone encodes a single *Rpi-blb* specific transcript. The coding sequence of the *Rpi-blb* transcript is 2913 nucleotides The putative *Rpi-blb* transcript is estimated to be 3138 nucleotides (nt) and contains a 44 and 181 nt long 5'- and 3'-untranslated region (UTR), respectively. The *Rpi-blb* gene contains a single intron of 678 nt starting 428 nt after the translational ATG start codon of the gene (Figure 3C).

The deduced open reading frame of the *Rpi-blb* gene encodes a predicted polypeptide of 970 amino acids with an estimated molecular weight of 110.3 kD (Figure 8). Several functional motifs present in *R* genes of the NBS-LRR class of plant *R* genes are apparent in the encoded protein which can be subdivided into 3 domains (A, B and C; Figure 8). The N-terminal part of the protein contains potential coiled-coil domains, heptad repeats in which the first and fourth residues are generally hydrophobic (domain A). Domain B harbours the NBS and other motifs that constitute the NB-ARC domain (ARC for Apaf-1, R protein, and CED-4) of R proteins and cell death regulators in animals (van der Biezen and Jones, 1998). This domain includes the Ap-ATPase motifs present in proteins of eukaryotic and prokaryotic origin (Aravind et al., 1999 Trends Biochem. Sci. 24, 47-53). The C-terminal half of Rpi-blb comprises a series of 19-20 irregular LRRs (domain C). The LRRs can be aligned according to the consensus sequence LxxLxxLxLxxC/N/SxxLxxLPxxa, where x designates any residue and "a" designates the positions of aliphatic amino acids, followed by a region of varying length. This repeat format approximates the consensus for cytoplasmic LRRs (Jones and Jones, 1997 Adv. Bot. Res.24, 89-167).

EXAMPLE 7: NATURAL HOMOLOGUES AND ARTIFICIAL VARIANTS OF THE *Rpi-blb* GENE

Natural homologues

BLASTN homology searches with the coding DNA sequence of the *Rpi-blb* gene identified a number of sequences with significant homology to short stretches of the *Rpi-blb* gene (Figure 9C). Nucleotides 549-1245 of the coding sequence of the *Rpi-blb* gene share 81- 90% sequence identity to partial NBS fragments from tomato (e.g. Q194, Q137, Q198 and Q199; Pan et al., 2000 Genetics. 155:309-22). These homologous sequences vary in length between 525 and 708 nucleotides and are PCR fragments which were identified by systematically scanning the tomato genome using (degenerate) primer pairs based on ubiquitous NBS motifs (Pan et al., 2000 Genetics. 155:309-22; Leister et al., 1996 Nat Genet. 14:421-429). Another region of the *Rpi-blb* gene which shares significant homology to a state of the art sequence comprises nucleotides 76-805 of the coding sequence. This 729 nt long sequence shares 91% sequence identity to an EST from potato (EMBL database accession no. BG890602; Figure 9C). The *Rpi-blb* gene sequence downstream of nucleotide 1245, comprising the LRR region, shares no significant homology to any state of the art sequence. BLASTX homology searches with the coding sequence of the *Rpi-blb* gene revealed that amino acid sequence homology with various state of the art genes does not exceed 36% sequence identity (Table 4). The best BLASTX score was obtained with an NBS-LRR gene derived from *Oryza sativa* (36.5% amino acid sequence identity). NBS-LRR genes sharing an overall sequence homology of 27-36% amino-acid sequence identity with *Rpi-blb* can be found among others in *Arabidopsis thaliana*, *Phaseolus vulgaris*, *Lycopersicon esculentum* (*Fusarium I2* gene cluster; Ori et al., 1997 Plant Cell, 9, 521-532; Simons et al, 1998 Plant Cell 10, 1055-1068), *Zea mays*, *Hordeum vulgare* and *Lactuca sativa*. Phylogenetic studies of the deduced amino acid sequences of *Rpi-blb*, *RGC1-blb*, *RGC3-blb*, *RGC4-blb* and those of the homologous state of the art genes (as defined by BLASTX) derived from diverse species, using the Neighbour-Joining method of Saitou and Nei (1987 Molecular Biology and Evolution 4, 406-425), shows that members of the *Rpi-blb* gene cluster can be placed in a separate branch (Figure 9).

Sequence comparisons of the four *RGCs* of the *Rpi-blb* gene cluster identified on 8005-8 BAC clone SPB4 show that sequence homology within the *Rpi-blb* gene cluster varies between 70% and 81% at the amino acid level. The deduced amino acid sequence of *Rpi-blb* shares the highest overall homology with *RGC3-blb* (81% amino acid sequence identity; Table 4). When the different domains are compared it is clear

that the N-terminal halves of the proteins (coiled-coil and NB-ARC domains) share a higher degree of homology (91% amino acid sequence identity) than the C-terminal halves of these proteins (LRRs; 71% amino acid sequence identity). The N-terminus of NBS-LRR proteins influences the requirement for downstream signalling components and is therefore thought to be the putative effector domain (Feys and Parker, 2000 Trends Genet 16:449-55). The C-terminal LRR region is implicated, by genetic studies, in elicitor recognition specificity (Ellis et al., 2000 Trends Plant Sci. 5:373-379; Dodds et al., 2001 Plant Cell 13:163-78).

Comparison of all four amino acid sequences revealed a total of 104 Rpi-blb specific amino acid residues (Figure 10A). The majority of these are located in the LRR region (80/104). Within the latter region, these specific residues are concentrated in the LRR subdomain xxLxLxxxx. The relative frequency of these specific amino-acid residues within this LRR subdomain is more than two times higher (28.3%) than that observed in the rest of the LRR domain (12.3%). The residues positioned around the two conserved leucine residues in the consensus xxLxxLxxxx are thought to be solvent exposed and are therefore likely to be involved in creating/maintaining recognition specificity of the resistance protein.

Sequences of additional homologues of the *Rpi-blb* gene can be obtained by screening genomic DNA or insert libraries, e.g. BAC libraries with primers based on signature sequences of the *Rpi-blb* gene. Screening of various *Solanum* BAC libraries with primer sets A and/or B (Table 2 and Figure 7) identified other *Rpi-blb* homologues derived from *Solanum bulbocastanum* (B149-blb), *S. tuberosum* (SH10-tub and SH20-tub) and *S. tarijense* (T118-tar). Comparison of all 8 protein sequences reduces the number of Rpi-blb specific amino acid residues to 51 (51/970; 5.25%) (Figure 10B). The majority of these are located in the LRR region (42/51; 82%). The relative frequency of these specific amino-acid residues within the LRR subdomain xxLxlxxxx is 3.3 times higher than that observed in the rest of the LRR domain (18.8% versus 5.7%, respectively). These data clearly suggest that evolution of *P. infestans* resistance specificity within the *Rpi-blb* gene cluster has mainly evolved through shifts in *Rpi-blb* LRR specific residues.

Inclusion of the additional *Rpi-blb* homologues in the above described phylogenetic tree analyses, using the Neighbour-Joining method of Saitou and Nei (1987 Molecular Biology and Evolution 4, 406-425), further justifies phylogenetic tree analysis as a method to define *Rpi-blb* homologous sequences (Figure 9B). Any functional R gene product which shares at least 70% sequence identity at the amino

acid level will end up in the same branch as gene products of the the *Rpi-blb* gene cluster and can thus be defined as being a homologue of *Rpi-blb*.

Artificial variants

Domain swaps between the different homologues can be made to ascertain the role of the different sequences in *P. infestans* resistance. The restriction enzyme *Nsi*I for example, which recognises the DNA sequence ATGCAT present in the conserved MHD motif can be used to swap the complete LRR domain of *Rpi-blb* with that of *RGC1-blb* or *RGC3-blb* using techniques known to those skilled in the art. Chimeric variants of the *Rpi-blb* gene were made which encode the N-terminal half of Rpi-blb and the C-terminal half of RGC1-blb or RGC3-blb and visa versa, i.e., the N-terminal half of RGC1-blb or RGC3-blb and the C-terminal half of Rpi-blb (Figure 11). These variants were transformed to the susceptible potato genotype Impala and tested for *P. infestans* resistance. Chimeric *RGC3-blb* genes containing the LRR domain of *Rpi-blb* were resistant to *P. infestans* indicating that the specificity of the *Rpi-blb* gene is encoded by this part of the gene.

Table 1. Overview of *P. infestans* susceptibility in different *S. bulbocastanum* accessions

S. bulbocastanum accession			#	#	#	%	
CGN	BGRC	PI	Plants	R	V	susceptibility	Cluster [a]
17692	8005	275193	11	10	1	9	A
	8006	275194	16	15	1	6	A
17693	8008	275198	19	18		0	B
17687	7997	243505	35	25	4	14	B
17688	7999	255518	19	19	0	0	C

[a] The letters a, b and c represent relative geographical origins depicted in Figure 1

Table 2. Overview of markers used for mapping *Rpi-blb*

Marker	Ori[a]	Sequence[b]	Annealing Temp (°C)	Restriction enzyme[c]
TG513	F	CGTAAACGCACCAAAAGCAG	58	a.s.
	R	GATTCAAGCCAGGAACCGAG		
TG330	F	CAGCTGCCACAGCTCAAGC	56	TaqI
	R	TACCTACATGTACAGTACTGC		
CT88	F	GGCAGAAGAGCTAGGAAGAG	57	MboI
	R	ATGGCGTGATACAATCCGAG		
	F	TTCAAGAGCTTGAAGACATAACA	60	a.s.
	R	ATGGCGTGATACAATCCGAG		
CT64	F	ACTAGAGGATAGATTCTTGG	56	CfoI
	R	CTGGATGCCTTTCTCTATGT		
B139R	F	GATCAGAAGTGCCTTGAACC	56	TaqI
	R	CAAGGAGCTTGGTCAGCAG		
SPB33L	F	ATTGCACAGGAGCAGATCTG	59	HinfI
	R	TGTAAGAGAGCAAGAGGCAC		
SPB42L	F	AGAGCAGTCTTGAAGGTTGG	58	CfoI
	R	GATGGTAACTAAGCCTCAGG		
B149R	F	GACAGATTTCTCATAAACCTGC	58	MseI / XbaI
	R	AATCGTGCATCACTAGAGCG		
RGC1-4	F	TGTGGAGTAAGAGAGGAAGG	62	SspI / MseI
	R	TCAGCTGAGCAGTGTGTGG		
A	F	ATGGCTGAAGCTTTCATTCAAGTTCTG	60	
	R	TCACACCGCTTGATCAGTTGTGGAC		
B	F	TRCATGAYCTMATCCATGATTTGC	60	
	R	GMAATTTTGTGCCAGTCTTCTCC		

[a] Orientation of the primer, F: forward, R: reverse

[b] primer sequences according to IUB codes

[c] a.s.: allele specific.

Table 3. Complementation of late blight susceptibility in potato and tomato

Genotype[a]	RGA-containing plants/ transformants	R plants / RGA-containing plants
IMP(RGC1-blb)	15/17[b]	0/15
	8/9[d]	0/8
IMP(RGC2-blb)	6/31[c]	6/6
	12/14[d]	9/12
IMP(RGC3-blb)	0/6[c]	-
	5/5[d]	0/5
IMP(RGC4-blb)	18/19[b]	0/18
	1/12[c]	0/1
IMP(vector)	8/8[b]	0/8
	9/10[d]	0/9
MM(RGC2-blb)	9/11[d]	7/9

[a] Primary transformants obtained from transformation of the susceptible potato and tomato genotypes Impala (IMP) and Moneymaker (MM), respectively, with T-DNA constructs containing the *Rpi-blb* gene candidates *RGC1-blb*, *RGC2-blb*, *RGC3-blb* or *RGC4-blb*. *Agrobacterium tumefaciens* strains AGL0[b], LBA4404[c], or UIA143[d] were used for transformation. Resistance was tested in detached leaf assays using the complex isolates IPO655-2A and IPO428-2.

Table 4. Comparison of nucleotide and amino acid sequence homology

		8005-8 BAC SPB4			Rice	Arabidopsis	Tomato
		RGC3-blb	*RGC1-blb*	*RGC4-blb*	RGC	RGC	I2C-1
Rpi-blb	nt[a]	88	84	81	-	-	-
	aa[a]	81	76	70	36	32	32
		N[b] C[b]	N C	N C			
		91 71	79 72	75 66			

[a] Percentage nucleotide (nt) and amino acid (aa) sequence identity.

[b] Separate comparisons were made for the N-terminal (N) and C-terminal (C) halves of the protein sequences. The border between the two halves is the conserved *Nsi*I restriction site in the DNA sequence (position 1417 of the *Rpi-blb* coding sequence).

Claims

1. An isolated or recombinant nucleic acid comprising a nucleic acid coding for the amino acid sequence of fig. 8 or a functional fragment or a homologue thereof.
2. A fragment according to claim 1 coding for the leucine rich repeat (LRR) fragment of the amino acid sequence of fig. 8.
3. A nucleic acid according to claim 1 or 2, said nucleic acid encoding a gene product that is capable of providing a member of the *Solanaceae* family with resistance against an oomycete pathogen, or a functional equivalent thereof.
4. A nucleic acid according to claim 3 wherein said member of the *Solanaceae* family comprises *S. tuberosum*.
5. A nucleic acid according to claim 3 where said resistance is race non-specific.
6. A nucleic acid according to claim 1 to 5 comprising a sequence as depicted in figure 6 for Rpi-blb or part thereof.
7. A nucleic acid according to claim 1 to 5 at least comprising a LRR domain.
8. A vector comprising a nucleic acid according to anyone of claims 1 to 7.
9. A host cell comprising a nucleic acid according to anyone of claims 1 to 7 or a vector according to claim 8.
10. A cell according to claim 9 comprising a plant cell.
11. A cell according to claim 10 wherein said plant comprises a member of the *Solanaceae* family.
12. A plant comprising a cell according to anyone of claims 9 to 11.
13. A part derived from a plant according to claim 12.
14. A part according to claim 13 wherein said tuber comprises a potato or said fruit comprises a tomato.
15. Progeny of a plant according to claim 12.
16. A proteinaceous substance encoded by a nucleic acid according to anyone of claims 1 to 7.
17. A proteinaceous substance comprising an amino acid sequence as depicted in figure 8 or a functional equivalent thereof.

18. A binding molecule directed at a nucleic acid according to anyone of claim 1 to 7.

19. A binding molecule according to claim 18 comprising a probe or primer.

20. A binding molecule according to claim 18 or 19 provided with a label.

21. A binding molecule according to claim 20 wherein said label comprises an excitable moiety.

22. Use of a nucleic acid according to anyone of claims 1 to 7 or a vector according to claim 8 or a cell according to anyone of claims 9 to 11 or a substance according to claim 16 or 17 or a binding molecule according to anyone of claims 18 to 21 in a method for providing a plant or its progeny with resistance against an oomycete infection.

23. Use according to claim 22 wherein said oomycete comprises *Phytophthora infestans*.

24. Use according to claim 22 or 23 wherein said plant comprises *S. tuberosum*.

25. A method for providing a plant or its progeny with at least partial resistance against an oomycete infection comprising providing said plant or part thereof with a gene or functional fragment thereof comprising a nucleic acid corresponding to one of a cluster of genes identifiable by phylogenetic tree analyses as corresponding to the *Rpi-blb, RCG1-blb, RCG3-blb* and *RCG4-blb* cluster of figure 9, said nucleic acid encoding a gene product that is capable of providing a member of the *Solanaceae* with resistance against an oomycete fungus, or providing said plant or part thereof with a nucleic acid according to anyone of claims 1 to 7 or a vector according to claim 8 or a cell according to claims 9-11 or a substance according to claim 16 or 17.

26. A method for selecting a plant or plant material or progeny thereof for its susceptibility or resistance to an oomycete infection comprising testing at least part of said plant or plant material or progeny thereof for the presence or absence of a nucleic acid corresponding to one of a cluster of genes identifiable by phylogenetic tree analyses as corresponding to the *Rpi-blb, RCG1-blb, RCG3-blb* and *RCG4-blb*cluster of figure 9, said nucleic acid encoding a gene product that is capable of providing a member of the *Solanaceae* with resistance against an oomycete fungus.

27. A method according to claim 26 comprising contacting at least part of said plant or plant material or progeny thereof with a binding molecule according to anyone of claims 18 to 21 and determining the binding of said molecule to said part.

28. A method according to claim 27 wherein said oomycete comprises *Phytophthora infestans*.

29. A method according to claim 27 or 28 wherein said plant comprises *S. tuberosum*.

30. An isolated *S. bulbocastanum*, or part thereof, susceptible to an oomycete infection caused by *Phytophthora infestans*.

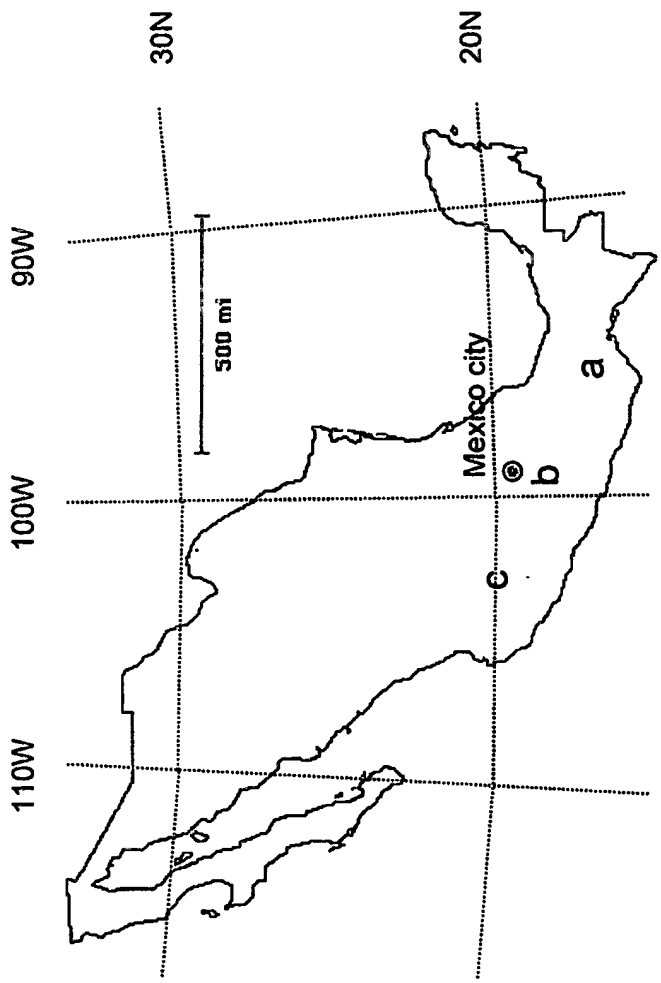

Figure 1

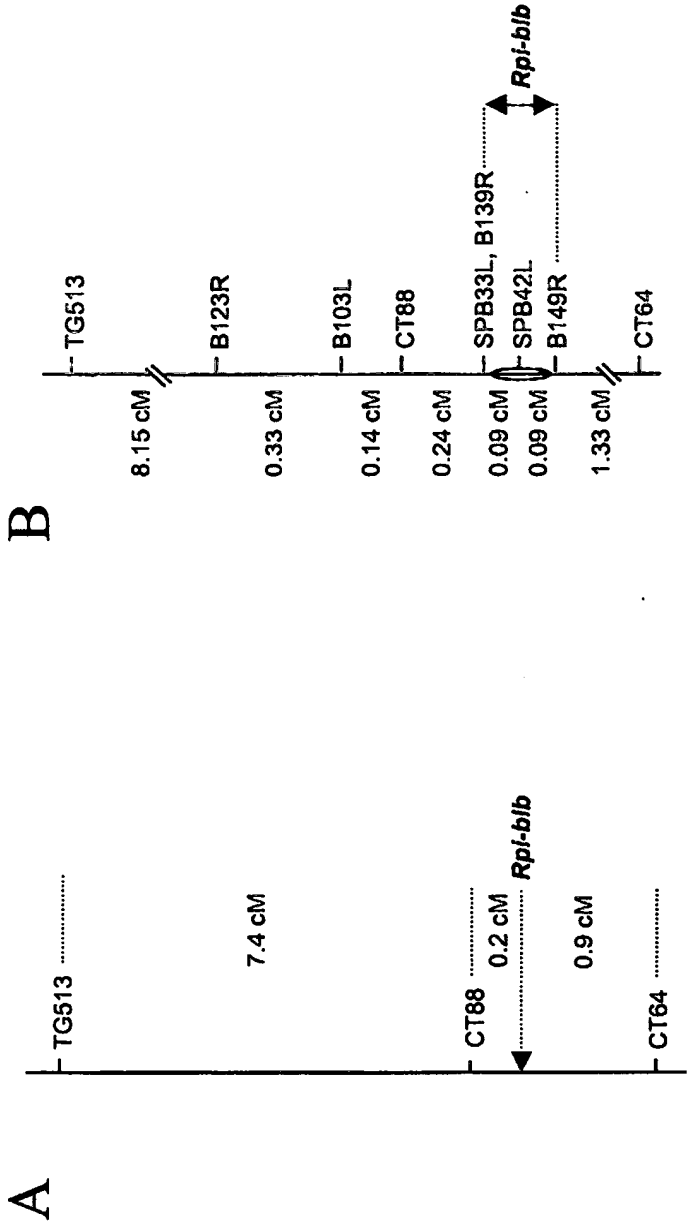
A
TG513
7.4 cM
CT88
0.2 cM
Rpi-blb
0.9 cM
CT64
B
TG513
8.15 cM
B123R
0.33 cM
B103L
0.14 cM
CT88
0.24 cM
SPB33L, B139R
0.09 cM
SPB42L
0.09 cM
B149R
1.33 cM
CT64
Rpi-blb

Figure 2

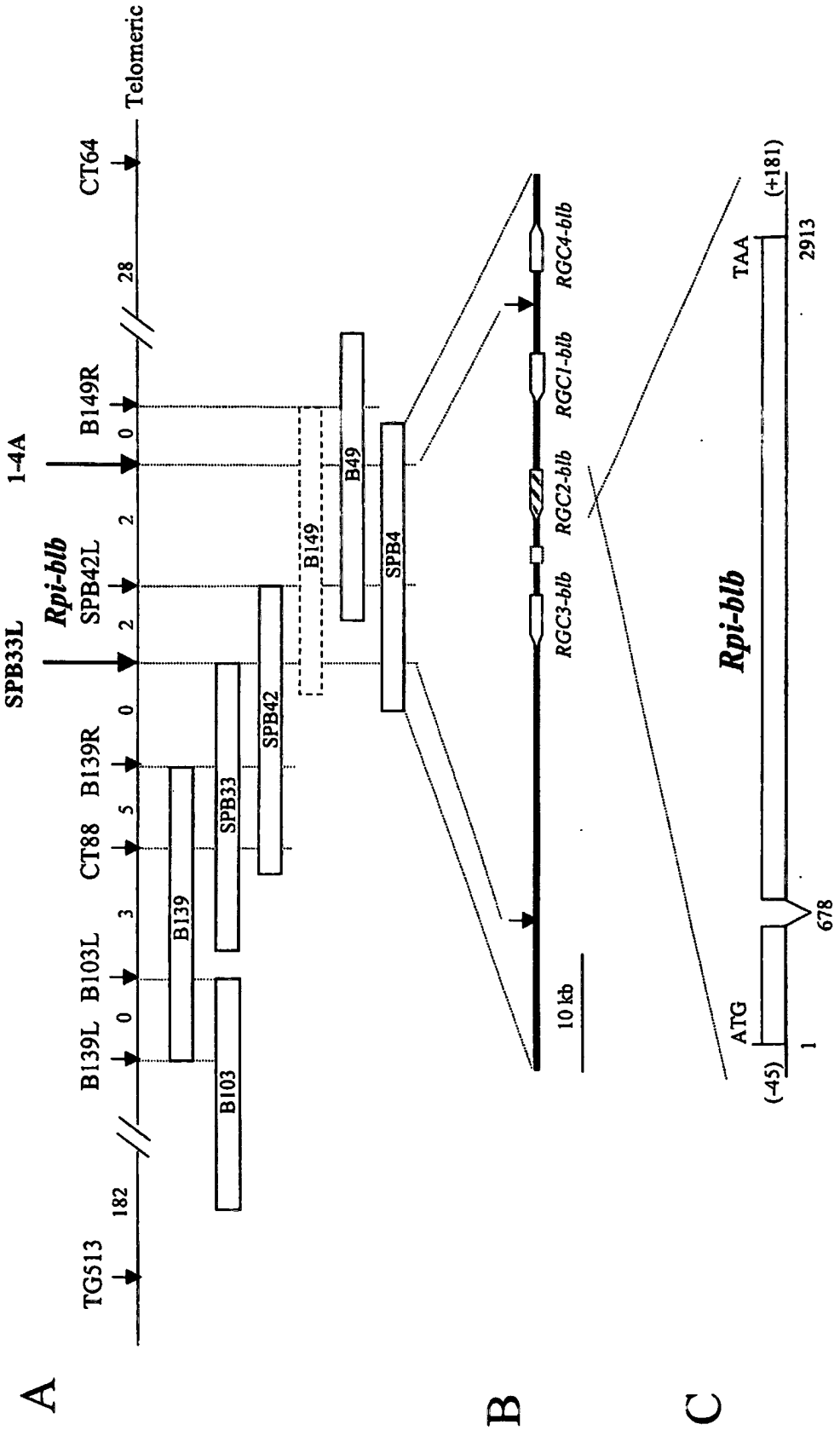
A
B
C
TG513
182
B139L
B103L
CT88
B139R
SPB33L
Rpi-blb
SPB42L
1-4A
B149R
CT64
Telomeric
B139
B103
SPB33
SPB42
B149
B49
SPB4
RGC3-blb
RGC2-blb
RGC1-blb
RGC4-blb
10 kb
Rpi-blb
ATG
TAA
(-45)
1
678
2913
(+181)

Figure 3

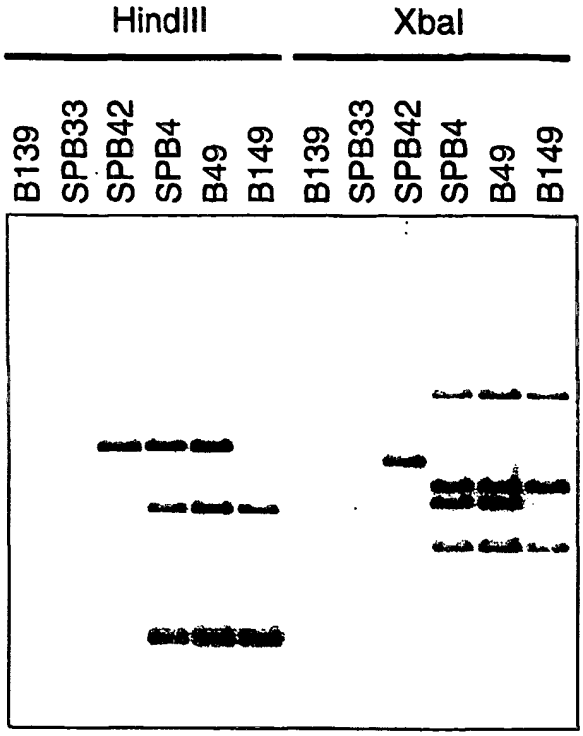
HindIII
XbaI
B139
SPB33
SPB42
SPB4
B49
B149
B139
SPB33
SPB42
SPB4
B49
B149

Figure 4

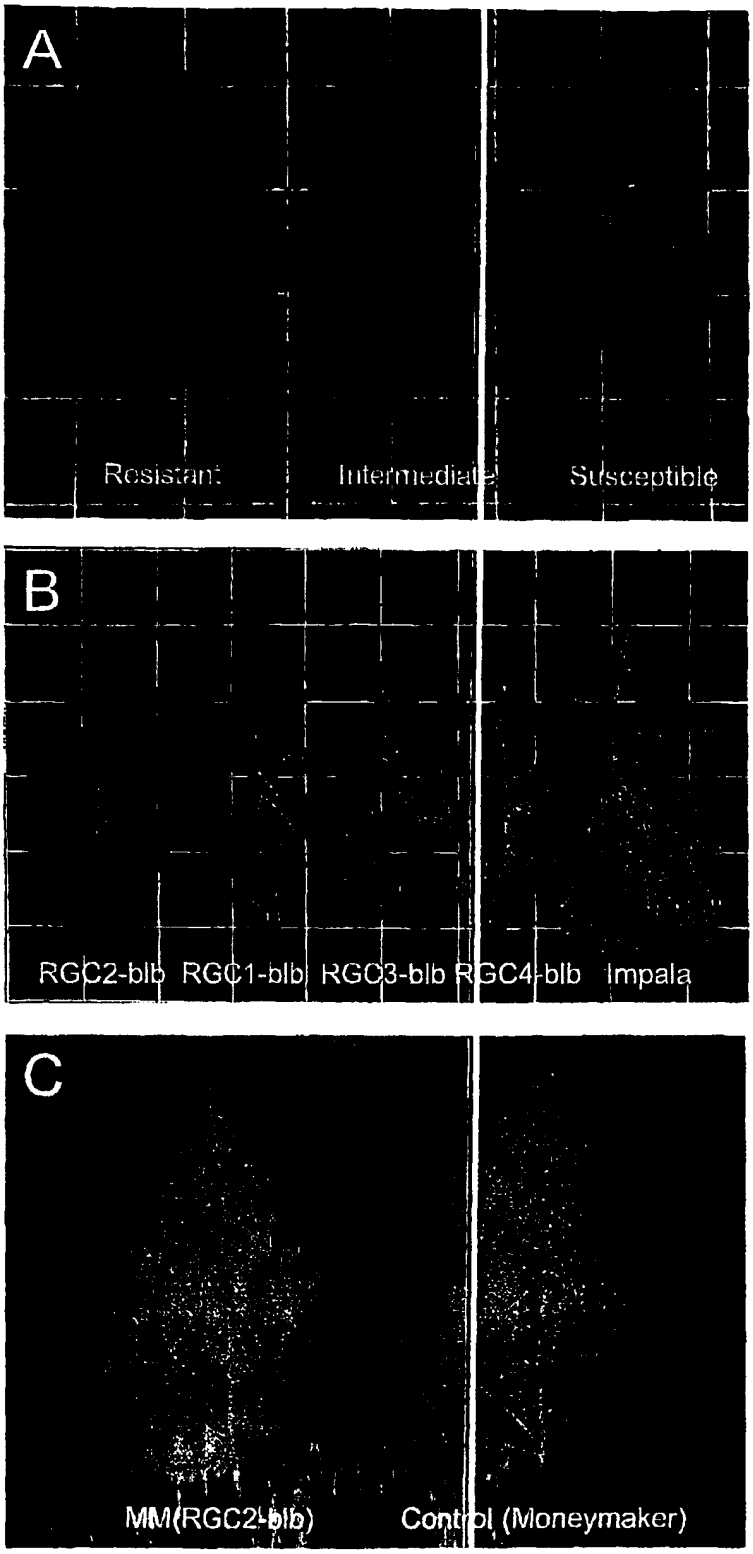
A
Resistant
Intermediate
Susceptible
B
RGC2-blb
RGC1-blb
RGC3-blb
RGC4-blb
Impala
C
MM(RGC2-blb)
Control (Moneymaker)

Figure 5

Figure 6A

```
1       ATGGCTGAAGCTTTCATTCAAGTTCTGCTAGACAATCTCACTTCTTTCCT
51      CAAAGGGGAACTTGTATTGCTTTTCGGTTTTCAAGATGAGTTCCAAAGGC
101     TTTCAAGCATGTTTTCTACAATTCAAGCCGTCCTTGAAGATGCTCAGGAG
151     AAGCAACTCAACAACAAGCCTCTAGAAAATTGGTTGCAAAAACTCAATGC
201     TGCTACATATGAAGTCGATGACATCTTGGATGAATATAAAACCAAGGCCA
251     CAAGATTCTCCCAGTCTGAATATGGCCGTTATCATCCAAAGGTTATCCCT
301     TTCCGTCACAAGGTCGGGAAAAGGATGGACCAAGTGATGAAAAAACTAAA
351     GGCAATTGCTGAGGAAAGAAAGAATTTTCATTTGCACGAAAAAATTGTAG
401     AGAGACAAGCTGTTAGACGGGAAACAGGTTCTGTATTAACCGAACCGCAG
451     GTTTATGGAAGAGACAAAGAGAAAGATGAGATAGTGAAAATCCTAATAAA
501     CAATGTTAGTGATGCCCAACACCTTTCAGTCCTCCCAATACTTGGTATGG
551     GGGGATTAGGAAAAACGACTCTTGCCCAAATGGTCTTCAATGACCAGAGA
601     GTTACTGAGCATTTCCATTCCAAAATATGGATTTGTGTCTCGGAAGATTT
651     TGATGAGAAGAGGTTAATAAAGGCAATTGTAGAATCTATTGAAGGAAGGC
701     CACTACTTGGTGAGATGGACTTGGCTCCACTTCAAAAGAAGCTTCAGGAG
751     TTGCTGAATGGAAAAAGATACTTGCTTGTCTTAGATGATGTTTGGAATGA
801     AGATCAACAGAAGTGGGCTAATTTAAGAGCAGTCTTGAAGGTTGGAGCAA
851     GTGGTGCTTCTGTTCTAACCACTACTCGTCTTGAAAAGGTTGGATCAATT
901     ATGGGAACATTGCAACCATATGAACTGTCAAATCTGTCTCAAGAAGATTG
951     TTGGTTGTTGTTCATGCAACGTGCATTTGGACACCAAGAAGAAATAAATC
1001    CAAACCTTGTGGCAATCGGAAAGGAGATTGTGAAAAAAAGTGGTGGTGTG
```

1051 CCTCTAGCAGCCAAAACTCTTGGAGGTATTTTGTGCTTCAAGAGAGAAGA

1101 AAGAGCATGGGAACATGTGAGAGACAGTCCGATTTGGAATTTGCCTCAAG

1151 ATGAAAGTTCTATTCTGCCTGCCCTGAGGCTTAGTTACCATCAACTTCCA

1201 CTTGATTTGAAACAATGCTTTGCGTATTGTGCGGTGTTCCCAAAGGATGC

1251 CAAAATGGAAAAAGAAAAGCTAATCTCTCTCTGGATGGCGCATGGTTTTC

1301 TTTTATCAAAAGGAAACATGGAGCTAGAGGATGTGGGCGATGAAGTATGG

1351 AAAGAATTATACTTGAGGTCTTTTTTCCAAGAGATTGAAGTTAAAGATGG

1401 TAAAACTTATTTCAAGATGCATGATCTCATCCATGATTTGGCAACATCTC

1451 TGTTTTCAGCAAACACATCAAGCAGCAATATCCGTGAAATAAATAAACAC

1501 AGTTACACACATATGATGTCCATTGGTTTCGCCGAAGTGGTGTTTTTTTA

1551 CACTCTTCCCCCCTTGGAAAAGTTTATCTCGTTAAGAGTGCTTAATCTAG

1601 GTGATTCGACATTTAATAAGTTACCATCTTCCATTGGAGATCTAGTACAT

1651 TTAAGATACTTGAACCTGTATGGCAGTGGCATGCGTAGTCTTCCAAAGCA

1701 GTTATGCAAGCTTCAAAATCTGCAAACTCTTGATCTACAATATTGCACCA

1751 AGCTTTGTTGTTTGCCAAAAGAAACAAGTAAACTTGGTAGTCTCCGAAAT

1801 CTTTTACTTGATGGTAGCCAGTCATTGACTTGTATGCCACCAAGGATAGG

1851 ATCATTGACATGCCTTAAGACTCTAGGTCAATTTGTTGTTGGAAGGAAGA

1901 AAGGTTATCAACTTGGTGAACTAGGAAACCTAAATCTCTATGGCTCAATT

1951 AAAATCTCGCATCTTGAGAGAGTGAAGAATGATAAGGACGCAAAAGAAGC

2001 CAATTTATCTGCAAAAGGGAATCTGCATTCTTTAAGCATGAGTTGGAATA

2051 ACTTTGGACCACATATATATGAATCAGAAGAAGTTAAAGTGCTTGAAGCC

2101 CTCAAACCACACTCCAATCTGACTTCTTTAAAAATCTATGGCTTCAGAGG

2151 AATCCATCTCCCAGAGTGGATGAATCACTCAGTATTGAAAAATATTGTCT

2201 CTATTCTAATTAGCAACTTCAGAAACTGCTCATGCTTACCACCCTTTGGT

2251 GATCTGCCTTGTCTAGAAAGTCTAGAGTTACACTGGGGGTCTGCGGATGT

2301 GGAGTATGTTGAAGAAGTGGATATTGATGTTCATTCTGGATTCCCCACAA

2351 GAATAAGGTTTCCATCCTTGAGGAAACTTGATATATGGGACTTTGGTAGT

2401 CTGAAAGGATTGCTGAAAAAGGAAGGAGAAGAGCAATTCCCTGTGCTTGA

2451 AGAGATGATAATTCACGAGTGCCCTTTTCTGACCCTTTCTTCTAATCTTA

2501 GGGCTCTTACTTCCCTCAGAATTTGCTATAATAAAGTAGCTACTTCATTC

2551 CCAGAAGAGATGTTCAAAAACCTTGCAAATCTCAAATACTTGACAATCTC

2601 TCGGTGCAATAATCTCAAAGAGCTGCCTACCAGCTTGGCTAGTCTGAATG

2651 CTTTGAAAAGTCTAAAAATTCAATTGTGTTGCGCACTAGAGAGTCTCCCT

2701 GAGGAAGGGCTGGAAGGTTTATCTTCACTCACAGAGTTATTTGTTGAACA

2751 CTGTAACATGCTAAAATGTTTACCAGAGGGATTGCAGCACCTAACAACCC

2801 TCACAAGTTTAAAAATTCGGGGATGTCCACAACTGATCAAGCGGTGTGAG

2851 AAGGGAATAGGAGAAGACTGGCACAAAATTTCTCACATTCCTAATGTGAA

2901 TATATATATTTAA

Figure 6B

1 ATGGCTGAAGCTTTCATTCAAGTTCTGCTAGACAATCTCACTTCTTTCCT
51 CAAAGGGGAACTTGTATTGCTTTTCGGTTTTCAAGATGAGTTCCAAAGGC
101 TTTCAAGCATGTTTTCTACAATTCAAGCCGTCCTTGAAGATGCTCAGGAG
151 AAGCAACTCAACAACAAGCCTCTAGAAAATTGGTTGCAAAAACTCAATGC
201 TGCTACATATGAAGTCGATGACATCTTGGATGAATATAAAACCAAGGCCA
251 CAAGATTCTCCCAGTCTGAATATGGCCGTTATCATCCAAAGGTTATCCCT
301 TTCCGTCACAAGGTCGGGAAAAGGATGGACCAAGTGATGAAAAAACTAAA
351 GGCAATTGCTGAGGAAAGAAAGAATTTTCATTTGCACGAAAAAATTGTAG
401 AGAGACAAGCTGTTAGACGGGAAACAG**GTACTCATCTTAAATTAGTATTA**
451 **CAACAACTAAGTTTATATTCATTTTTTTGGCAATTATCAAATTCAGAAAA**
501 **GGGTTAAATATACTCATGTCCTATCGTAAATAGTGTATATATACCTCTCG**
551 **TTGTACTTTCGATCTGAATATACTTGTCAAATCTGGCAAGCTCAGAATCA**
601 **AATTATCCACCCCAACTTTTAAATACTCGATATCTTTAGAAATCCACCTG**
651 **TCTAACTCATCCACTACCCATTCCCTTTGCTTTGAATTCTTTTCTTTACC**
701 **TATAAACTTGGAACACTCGATCCGTTTTGCTTTTCTTAACAAAGCAGCTC**
751 **AGAGAAAAGAGGTTTTCTTCTATTCTGTTTCTCTGTGTGCTGCACTTGGG**
801 **TCCTTAATCCCATTAAAAACAGGGCATGTTAATCCCAACGACGGTAGCCT**
851 **TTCCTGACAGCTGACTGTAAATTTTGTCTAACAAAGAAAAAAAAAGATTA**
901 **GACATGTTTTTCCTTGTCATTGATTAGGCTGGATTTCTTTCAGAGTGGAA**
951 **CATAGGGGATATATTGGACCAAAAGTAGAATGGGTATATATTTAAAGTAT**
1001 **TTCTGATAGAACAGGAGTATATTGTGCGAAAATATCCTCTATTTTCTGTT**
1051 **GTCTCCTAATGAGTTTGAATGTAATAATATTCTCATGTGGACATTGCTTG**
1101 **CACCAG**GTTCTGTATTAACCGAACCGCAGGTTTATGGAAGAGACAAAGAG
1151 AAAGATGAGATAGTGAAAATCCTAATAAACAATGTTAGTGATGCCCAACA
1201 CCTTTCAGTCCTCCCAATACTTGGTATGGGGGGATTAGGAAAAACGACTC
1251 TTGCCCAAATGGTCTTCAATGACCAGAGAGTTACTGAGCATTTCCATTCC
1301 AAAATATGGATTTGTGTCTCGGAAGATTTTGATGAGAAGAGGTTAATAAA
1351 GGCAATTGTAGAATCTATTGAAGGAAGGCCACTACTTGGTGAGATGGACT
1401 TGGCTCCACTTCAAAAGAAGCTTCAGGAGTTGCTGAATGGAAAAAGATAC

1451 TTGCTTGTCTTAGATGATGTTTGGAATGAAGATCAACAGAAGTGGGCTAA
1501 TTTAAGAGCAGTCTTGAAGGTTGGAGCAAGTGGTGCTTCTGTTCTAACCA
1551 CTACTCGTCTTGAAAAGGTTGGATCAATTATGGGAACATTGCAACCATAT
1601 GAACTGTCAAATCTGTCTCAAGAAGATTGTTGGTTGTTGTTCATGCAACG
1651 TGCATTTGGACACCAAGAAGAAATAAATCCAAACCTTGTGGCAATCGGAA
1701 AGGAGATTGTGAAAAAAAGTGGTGGTGTGCCTCTAGCAGCCAAAACTCTT
1751 GGAGGTATTTTGTGCTTCAAGAGAGAAGAAAGAGCATGGGAACATGTGAG
1801 AGACAGTCCGATTTGGAATTTGCCTCAAGATGAAAGTTCTATTCTGCCTG
1851 CCCTGAGGCTTAGTTACCATCAACTTCCACTTGATTTGAAACAATGCTTT
1901 GCGTATTGTGCGGTGTTCCCAAAGGATGCCAAAATGGAAAAAGAAAAGCT
1951 AATCTCTCTCTGGATGGCGCATGGTTTTCTTTTATCAAAAGGAAACATGG
2001 AGCTAGAGGATGTGGGCGATGAAGTATGGAAAGAATTATACTTGAGGTCT
2051 TTTTTCCAAGAGATTGAAGTTAAAGATGGTAAAACTTATTTCAAGATGCA
2101 TGATCTCATCCATGATTTGGCAACATCTCTGTTTTCAGCAAACACATCAA
2151 GCAGCAATATCCGTGAAATAAATAAACACAGTTACACACATATGATGTCC
2201 ATTGGTTTCGCCGAAGTGGTGTTTTTTTACACTCTTCCCCCCTTGGAAAA
2251 GTTTATCTCGTTAAGAGTGCTTAATCTAGGTGATTCGACATTTAATAAGT
2301 TACCATCTTCCATTGGAGATCTAGTACATTTAAGATACTTGAACCTGTAT
2351 GGCAGTGGCATGCGTAGTCTTCCAAAGCAGTTATGCAAGCTTCAAAATCT
2401 GCAAACTCTTGATCTACAATATTGCACCAAGCTTTGTTGTTTGCCAAAAG
2451 AAACAAGTAAACTTGGTAGTCTCCGAAATCTTTTACTTGATGGTAGCCAG
2501 TCATTGACTTGTATGCCACCAAGGATAGGATCATTGACATGCCTTAAGAC
2551 TCTAGGTCAATTTGTTGTTGGAAGGAAGAAAGGTTATCAACTTGGTGAAC
2601 TAGGAAACCTAAATCTCTATGGCTCAATTAAAATCTCGCATCTTGAGAGA
2651 GTGAAGAATGATAAGGACGCAAAAGAAGCCAATTTATCTGCAAAAGGGAA
2701 TCTGCATTCTTTAAGCATGAGTTGGAATAACTTTGGACCACATATATATG
2751 AATCAGAAGAAGTTAAAGTGCTTGAAGCCCTCAAACCACACTCCAATCTG
2801 ACTTCTTTAAAAATCTATGGCTTCAGAGGAATCCATCTCCCAGAGTGGAT
2851 GAATCACTCAGTATTGAAAAATATTGTCTCTATTCTAATTAGCAACTTCA
2901 GAAACTGCTCATGCTTACCACCCTTTGGTGATCTGCCTTGTCTAGAAAGT
2951 CTAGAGTTACACTGGGGGTCTGCGGATGTGGAGTATGTTGAAGAAGTGGA
3001 TATTGATGTTCATTCTGGATTCCCCACAAGAATAAGGTTTCCATCCTTGA

3051 GGAAACTTGATATATGGGACTTTGGTAGTCTGAAAGGATTGCTGAAAAAG
3101 GAAGGAGAAGAGCAATTCCCTGTGCTTGAAGAGATGATAATTCACGAGTG
3151 CCCTTTTCTGACCCTTTCTTCTAATCTTAGGGCTCTTACTTCCCTCAGAA
3201 TTTGCTATAATAAAGTAGCTACTTCATTCCCAGAAGAGATGTTCAAAAAC
3251 CTTGCAAATCTCAAATACTTGACAATCTCTCGGTGCAATAATCTCAAAGA
3301 GCTGCCTACCAGCTTGGCTAGTCTGAATGCTTTGAAAAGTCTAAAAATTC
3351 AATTGTGTTGCGCACTAGAGAGTCTCCCTGAGGAAGGGCTGGAAGGTTTA
3401 TCTTCACTCACAGAGTTATTTGTTGAACACTGTAACATGCTAAAATGTTT
3451 ACCAGAGGGATTGCAGCACCTAACAACCCTCACAAGTTTAAAAATTCGGG
3501 GATGTCCACAACTGATCAAGCGGTGTGAGAAGGGAATAGGAGAAGACTGG
3551 CACAAAATTTCTCACATTCCTAATGTGAATATATATATTTAA

Figure 6C

```
1     AGTACTCCATCCGTTCACTTTGATTTGTCATGTTGCACTTTTCGAAAGTC
51    AATTTGACTAATTTTTAAAGCTAAATTAGATTACACTAATTCAATATTTT
101   AAACAGAAAAATTAGATATTCAAAAACTATACAAAAAATATTATACATTG
151   CAATTTTTTGCATATCAATATGATAAAAAAATATATCGTAAAATATTAGT
201   CAAAATTTTTATAATTTGACTCAAATCATGAAAAGTATAATAATTAATAG
251   TGGACGGAGGAAGTATTGTCTTTCCAGATTTGTGGCCATTTTTGGTCCAA
301   GGGCCATTAGCAGTTCTCTTCATTTTCTACTTCTGTCTCATATTAGATGG
351   GCATCTTACTAAAAATATTTGTCTCATATTACTTGATTATTTATTAAATC
401   AAAAAGAATTAATTAATTTTTTCTCATTTTACCCCTACAATTAATATAGT
451   TTTAAAAGTTTTAAACAAATTTTGAAGAATCAAAATTTCTTTTGCAAGAG
501   ACTTATTAATATAAACAAAGGATAAAATAATAAAAGCTGTCAATTTATTG
551   ACCATCACTTAATAATATATAAAATACAAACTGCTGATCTAATATGAGAC
601   GGACAAAATATATTCTAAAATATTTTCGGACAGATATGTGATATTCTAAC
651   CATTCACTACACTATATTATGCATTTTATCCGCCAATGACTTATTTCAGC
701   TTTAATTAATTAGGAAAGAGGAAACTGCCAATGAGGAAGAGTAGGGGCGT
751   AGTTGCTGTCGACGAAAAAAAGATAATACTCACTCTTTTCGATTTTTATT
801   TTTATTTATCACTTTTAACCTATCATGTAAAAAGATAATTATTTTTTTCA
851   TGCTTTATCCTTAGTATTAAACAATTTAATAGGGATTATTTTGTAAAATA
901   TTTATATGAATAATTGTTTTCGTAATGAATTTGTCCGGTCAAACAATGAT
951   AAATAAAAATGAATGAAGAGAGTAGAAAACAAAACAAAAGAACAAGTTGA
1001  CAACTTGAGAGATTAAAAGGGTCCAAAACGCCTTGGATTTTGAGATTCCA
1051  TATGTGAAATTTCCATGAAATAATTGAATTTGTATTATTACAAGTCAAAC
1101  TTTCCATTTCATTCCAACTAGCCATCTTGGTTTCAAAATTACACATTCAT
1151  TCATTCACAGATCTAATATTCTTAATAGTGATTTCCACATATGGCTGAAG
1201  CTTTCATTCAAGTTCTGCTAGACAATCTCACTTCTTTCCTCAAAGGGGAA
1251  CTTGTATTGCTTTTCGGTTTTCAAGATGAGTTCCAAAGGCTTTCAAGCAT
1301  GTTTTCTACAATTCAAGCCGTCCTTGAAGATGCTCAGGAGAAGCAACTCA
1351  ACAACAAGCCTCTAGAAAATTGGTTGCAAAAACTCAATGCTGCTACATAT
1401  GAAGTCGATGACATCTTGGATGAATATAAAACCAAGGCCACAAGATTCTC
```

1451 CCAGTCTGAATATGGCCGTTATCATCCAAAGGTTATCCCTTTCCGTCACA
1501 AGGTCGGGAAAAGGATGGACCAAGTGATGAAAAAACTAAAGGCAATTGCT
1551 GAGGAAAGAAAGAATTTTCATTTGCACGAAAAAATTGTAGAGAGACAAGC
1601 TGTTAGACGGGAAACAGGTACTCATCTTAAATTAGTATTACAACAACTAA
1651 GTTTATATTCATTTTTTTGGCAATTATCAAATTCAGAAAAGGGTTAAATA
1701 TACTCATGTCCTATCGTAAATAGTGTATATATACCTCTCGTTGTACTTTC
1751 GATCTGAATATACTTGTCAAATCTGGCAAGCTCAGAATCAAATTATCCAC
1801 CCCAACTTTTAAATACTCGATATCTTTAGAAATCCACCTGTCTAACTCAT
1851 CCACTACCCATTCCCTTTGCTTTGAATTCTTTTCTTTACCTATAAACTTG
1901 GAACACTCGATCCGTTTTGCTTTTCTTAACAAAGCAGCTCAGAGAAAAGA
1951 GGTTTTCTTCTATTCTGTTTCTCTGTGTGCTGCACTTGGGTCCTTAATCC
2001 CATTAAAAACAGGGCATGTTAATCCCAACGACGGTAGCCTTTCCTGACAG
2051 CTGACTGTAAATTTTGTCTAACAAAGAAAAAAAAAGATTAGACATGTTTT
2101 TCCTTGTCATTGATTAGGCTGGATTTCTTTCAGAGTGGAACATAGGGGAT
2151 ATATTGGACCAAAAGTAGAATGGGTATATATTTAAAGTATTTCTGATAGA
2201 ACAGGAGTATATTGTGCGAAAATATCCTCTATTTTCTGTTGTCTCCTAAT
2251 GAGTTTGAATGTAATAATATTCTCATGTGGACATTGCTTGCACCAGGTTC
2301 TGTATTAACCGAACCGCAGGTTTATGGAAGAGACAAAGAGAAAGATGAGA
2351 TAGTGAAAATCCTAATAAACAATGTTAGTGATGCCCAACACCTTTCAGTC
2401 CTCCCAATACTTGGTATGGGGGGATTAGGAAAAACGACTCTTGCCCAAAT
2451 GGTCTTCAATGACCAGAGAGTTACTGAGCATTTCCATTCCAAAATATGGA
2501 TTTGTGTCTCGGAAGATTTTGATGAGAAGAGGTTAATAAAGGCAATTGTA
2551 GAATCTATTGAAGGAAGGCCACTACTTGGTGAGATGGACTTGGCTCCACT
2601 TCAAAAGAAGCTTCAGGAGTTGCTGAATGGAAAAAGATACTTGCTTGTCT
2651 TAGATGATGTTTGGAATGAAGATCAACAGAAGTGGGCTAATTTAAGAGCA
2701 GTCTTGAAGGTTGGAGCAAGTGGTGCTTCTGTTCTAACCACTACTCGTCT
2751 TGAAAAGGTTGGATCAATTATGGGAACATTGCAACCATATGAACTGTCAA
2801 ATCTGTCTCAAGAAGATTGTTGGTTGTTGTTCATGCAACGTGCATTTGGA
2851 CACCAAGAAGAAATAAATCCAAACCTTGTGGCAATCGGAAAGGAGATTGT
2901 GAAAAAAAGTGGTGGTGTGCCTCTAGCAGCCAAAACTCTTGGAGGTATTT
2951 TGTGCTTCAAGAGAGAAGAAAGAGCATGGGAACATGTGAGAGACAGTCCG
3001 ATTTGGAATTTGCCTCAAGATGAAAGTTCTATTCTGCCTGCCCTGAGGCT

3051 TAGTTACCATCAACTTCCACTTGATTTGAAACAATGCTTTGCGTATTGTG
3101 CGGTGTTCCCAAAGGATGCCAAAATGGAAAAAGAAAAGCTAATCTCTCTC
3151 TGGATGGCGCATGGTTTTCTTTTATCAAAAGGAAACATGGAGCTAGAGGA
3201 TGTGGGCGATGAAGTATGGAAAGAATTATACTTGAGGTCTTTTTTCCAAG
3251 AGATTGAAGTTAAAGATGGTAAAACTTATTTCAAGATGCATGATCTCATC
3301 CATGATTTGGCAACATCTCTGTTTTCAGCAAACACATCAAGCAGCAATAT
3351 CCGTGAAATAAATAAACACAGTTACACACATATGATGTCCATTGGTTTCG
3401 CCGAAGTGGTGTTTTTTTACACTCTTCCCCCCTTGGAAAAGTTTATCTCG
3451 TTAAGAGTGCTTAATCTAGGTGATTCGACATTTAATAAGTTACCATCTTC
3501 CATTGGAGATCTAGTACATTTAAGATACTTGAACCTGTATGGCAGTGGCA
3551 TGCGTAGTCTTCCAAAGCAGTTATGCAAGCTTCAAAATCTGCAAACTCTT
3601 GATCTACAATATTGCACCAAGCTTTGTTGTTTGCCAAAAGAAACAAGTAA
3651 ACTTGGTAGTCTCCGAAATCTTTTACTTGATGGTAGCCAGTCATTGACTT
3701 GTATGCCACCAAGGATAGGATCATTGACATGCCTTAAGACTCTAGGTCAA
3751 TTTGTTGTTGGAAGGAAGAAAGGTTATCAACTTGGTGAACTAGGAAACCT
3801 AAATCTCTATGGCTCAATTAAAATCTCGCATCTTGAGAGAGTGAAGAATG
3851 ATAAGGACGCAAAAGAAGCCAATTTATCTGCAAAAGGGAATCTGCATTCT
3901 TTAAGCATGAGTTGGAATAACTTTGGACCACATATATATGAATCAGAAGA
3951 AGTTAAAGTGCTTGAAGCCCTCAAACCACACTCCAATCTGACTTCTTTAA
4001 AAATCTATGGCTTCAGAGGAATCCATCTCCCAGAGTGGATGAATCACTCA
4051 GTATTGAAAAATATTGTCTCTATTCTAATTAGCAACTTCAGAAACTGCTC
4101 ATGCTTACCACCCTTTGGTGATCTGCCTTGTCTAGAAAGTCTAGAGTTAC
4151 ACTGGGGGTCTGCGGATGTGGAGTATGTTGAAGAAGTGGATATTGATGTT
4201 CATTCTGGATTCCCCACAAGAATAAGGTTTCCATCCTTGAGGAAACTTGA
4251 TATATGGGACTTTGGTAGTCTGAAAGGATTGCTGAAAAAGGAAGGAGAAG
4301 AGCAATTCCCTGTGCTTGAAGAGATGATAATTCACGAGTGCCCTTTTCTG
4351 ACCCTTTCTTCTAATCTTAGGGCTCTTACTTCCCTCAGAATTTGCTATAA
4401 TAAAGTAGCTACTTCATTCCCAGAAGAGATGTTCAAAAACCTTGCAAATC
4451 TCAAATACTTGACAATCTCTCGGTGCAATAATCTCAAAGAGCTGCCTACC
4501 AGCTTGGCTAGTCTGAATGCTTTGAAAAGTCTAAAAATTCAATTGTGTTG
4551 CGCACTAGAGAGTCTCCCTGAGGAAGGGCTGGAAGGTTTATCTTCACTCA
4601 CAGAGTTATTTGTTGAACACTGTAACATGCTAAAATGTTTACCAGAGGGA

4651 TTGCAGCACCTAACAACCCTCACAAGTTTAAAAATTCGGGGATGTCCACA
4701 ACTGATCAAGCGGTGTGAGAAGGGAATAGGAGAAGACTGGCACAAAATTT
4751 CTCACATTCCTAATGTGAATATATATATTTAAGTTATTTGCTATTGTTTC
4801 TTTGTTTGTGAGTCTTTTTGGTTCCTGCCATTGTGATTGCATGTAATTTT
4851 TTTCTAGGGTTGTTTCTTTATGAGTCTCTCTCTCATTGGATGTAATTTTC
4901 TTTTGGAAACAAATCTGTCAATTGATTTGTATTATACGCTTTCAGAATCT
4951 ATTACTTATTTGTAATTGTTTCTTTGTTTGTAAATTGTGAGTATCTTATT
5001 TTATGGAATTTTCTGATTTTATTTTGAAAACAAATCAATGATTTGTAAGA
5051 TCCATCTGTATTATACTCCCTTCGTCTCATTTTATGTGTCACCTGTCGGA
5101 TTTCGAGATTCAAACAAATCTATCTTTGATCGTAAATTTTTAATAGATCT
5151 TTTAAACATTTTGAATTATCAATTATTGTGACTTTAGTACT

Figure 6D

1 ATGGCTGAAGCTTTCCTTCAAGTTCTGCTAGATAATCTCACTTTTTTCAT
51 CCAAGGGGAACTTGGATTGGTTTTTGGTTTCGAGAAGGAGTTTAAAAAAC
101 TTTCAAGTATGTTTTCAATGATCCAAGCTGTGCTAGAAGATGCTCAAGAG
151 AAGCAACTGAAGTACAAGGCAATAAAGAACTGGTTACAGAAACTCAATGT
201 TGCTGCATATGAAGTTGATGACATCTTGGATGACTGTAAAACTGAGGCAG
251 CAAGATTCAAGCAGGCTGTATTGGGGCGTTATCATCCACGGACCATCACT
301 TTCTGTTACAAGGTGGGAAAAAGAATGAAAGAAATGATGGAAAAACTAGA
351 TGCAATTGCAGAGGAACGGAGGAATTTTCATTTAGATGAAAGGATTATAG
401 AGAGACAAGCTGCTAGACGGCAAACAG**GTGCTCATCTTAATTTTATTTTA**
451 **AAACAAATAAGTATTACAAATTGCAGAGAAACGAAGGAATTTATATTCAT**
501 **TTTTATTTTTGGCAATTATCAAAGTCATTTGTGTTTTTAAGCTGGGGGGA**
551 **AGTTTCAAATATTTTCTCTAGTCTTAATGTTTGTCTCACTCACTCAGCAT**
601 **GATTTTCTCAATCCTTCACTTCAACTCCCCCCTACTGTGCAAATATCTTC**
651 **TCTATTTTCTGTTGACTCCTAATGAGCTTGAATGTAACAACATTCTTGTT**
701 **TGGAGCAG**GTTTTGTTTTAACTGAGCCAAAAGTTTATGGAAGGGAAAAAG
751 AGGAGGATGAGATAGTGAAAATCTTGATAAACAATGTTAGTTATTCCGAA
801 GAAGTTCCAGTACTCCCAATACTTGGTATGGGGGGACTAGGAAAGACGAC
851 TCTAGCCCAAATGGTCTTCAATGATCAAAGAATTACTGAGCATTTCAATC
901 TAAAGATATGGGTTTGTGTCTCAGATGATTTTGATGAGAAGAGGTTGATT
951 AAGGCAATTGTAGAATCTATTGAAGGAAAGTCACTGGGTGACATGGACTT
1001 GGCTCCCCTCCAGAAAAAGCTTCAGGAGTTGTTGAATGGAAAAAGATACT
1051 TTCTTGTTTTGGATGATGTTTGGAATGAAGATCAAGAAAAGTGGGATAAT
1101 CTTAGAGCAGTATTGAAGATTGGAGCTAGTGGTGCTTCAATTCTAATTAC
1151 TACTCGTCTTGAAAAAATTGGATCAATTATGGGAACTTTGCAACTATATC
1201 AGTTATCAAATTTGTCTCAAGAAGATTGTTGGTTGTTGTTCAAGCAACGT
1251 GCATTTTGCCACCAAACCGAAACAAGTCCTAAACTTATGGAAATCGGAAA
1301 GGAGATTGTGAAGAAATGTGGGGGTGTGCCTCTAGCAGCCAAAACTCTTG
1351 GAGGCCTTTTACGCTTCAAGAGGGAAGAAAGTGAATGGGAACATGTGAGA
1401 GATAGTGAGATTTGGAATTTACCTCAAGATGAAAATTCTGTTTTGCCTGC

1451 CCTGAGGCTGAGTTATCATCATCTTCCACTTGATTTGAGACAATGTTTTG
1501 CATATTGCGCAGTATTCCCAAAGGACACCAAAATAGAAAAGGAATATCTC
1551 ATCGCTCTCTGGATGGCACACAGTTTTCTTTTATCAAAAGGAAACATGGA
1601 GCTAGAGGATGTGGGCAATGAAGTATGGAATGAATTATACTTGAGGTCTT
1651 TTTTCCAAGAGATTGAAGTTAAATCTGGTAAAACTTATTTCAAGATGCAT
1701 GATCTCATCCATGATTTGGCTACATCTATGTTTTCAGCAAGCGCATCAAG
1751 CAGAAGTATACGCCAAATAAATGTAAAAGATGATGAAGATATGATGTTCA
1801 TTGTAACAAATTATAAAGATATGATGTCCATTGGTTTCTCCGAAGTGGTG
1851 TCTTCTTACTCTCCTTCGCTCTTTAAAAGGTTTGTCTCGTTAAGGGTGCT
1901 TAATCTAAGTAACTCAGAATTTGAACAGTTACCGTCTTCCGTTGGAGATC
1951 TAGTACATTTAAGATACCTTGACCTGTCTGGTAATAAAATTTGTAGTCTT
2001 CCAAAGAGGTTGTGCAAGCTTCAAAATCTGCAGACTCTTGATCTATATAA
2051 TTGCCAGTCACTTTCTTGTTTGCCGAAACAAACAAGTAAGCTTTGTAGTC
2101 TCCGGAATCTTGTACTTGATCACTGTCCATTGACTTCTATGCCACCAAGA
2151 ATAGGATTGTTGACATGCCTTAAGACACTAGGTTACTTTGTTGTAGGCGA
2201 GAGGAAAGGTTATCAACTTGGTGAACTACGAAATTTAAACCTCCGTGGTG
2251 CAATTTCAATCACACATCTTGAGAGAGTGAAAAATGATATGGAGGCAAAA
2301 GAAGCCAATTTATCTGCAAAAGCAAATCTACACTCTTTAAGCATGAGTTG
2351 GGATAGACCAAACAGATATGAATCCGAAGAAGTTAAAGTGCTTGAAGCCC
2401 TCAAACCACATCCCAATCTGAAATATTTAGAAATCATTGACTTCTGTGGA
2451 TTCTGTCTCCCTGACTGGATGAATCACTCAGTTTTGAAAAATGTTGTCTC
2501 TATTCTAATTAGCGGTTGTGAAAACTGCTCGTGCTTACCACCCTTTGGTG
2551 AGCTGCCTTGTCTAGAAAGTCTGGAGTTACAAGACGGGTCTGTGGAGGTG
2601 GAGTATGTTGAAGATTCTGGATTCCTGACAAGAAGAAGATTTCCATCCCT
2651 GAGAAAACTTCATATAGGTGGCTTTTGTAATCTGAAAGGATTGCAGAGAA
2701 TGAAAGGAGCAGAGCAATTCCCCGTGCTTGAAGAGATGAAGATTTCGGAT
2751 TGCCCTATGTTTGTTTTTCCGACCCTTTCTTCTGTCAAGAAATTAGAAAT
2801 TTGGGGGGAGGCAGATGCAGGAGGTTTGAGCTCCATATCTAATCTCAGCA
2851 CTCTTACATCCCTCAAGATTTTCAGTAACCACACAGTGACTTCACTACTG
2901 GAAGAGATGTTCAAAAACCTTGAAAATCTCATATACTTGAGTGTCTCTTT
2951 CTTGGAGAATCTCAAAGAGCTGCCTACCAGCCTGGCTAGTCTCAACAATT
3001 TGAAGTGTCTGGATATTCGTTATTGTTACGCACTAGAGAGTCTCCCCGAG

3051 GAAGGGCTGGAAGGTTTATCTTCACTCACAGAGTTATTTGTTGAACACTG
3101 TAACATGCTAAAATGTTTACCAGAGGGATTGCAGCACCTAACAACCCTCA
3151 CAAGTTTAAAAATTCGGGGATGTCCACAACTGATCAAGCGGTGTGAGAAG
3201 GGAATAGGAGAAGACTGGCACAAAATTTCTCACATTCCTAATGTGAATAT
3251 ATATATTTAA

Figure 6E

1 ATGGCTGAAGCTTTCATTCAAGTTGTGCTAGACAATCTCACTTCTTTCCT
51 CAAAGGGGAACTTGTATTGCTTTTCGGTTTTCAAGATGAGTTCCAAAGGC
101 TTTCAAGCATGTTTTCTACAATCCAAGCCGTCCTTGAAGATGCTCAAGAG
151 AAGCAACTCAACGACAAGCCTCTAGAAAATTGGTTGCAAAAACTCAATGC
201 TGCTACATATGAAGTCGATGACATCTTGGATGAATATAAAACTAAGGCCA
251 CAAGATTCTTGCAGTCTGAATATGGCCGTTATCATCCAAAGGTTATCCCT
301 TTCCGTCACAAGGTTGGGAAAAGGATGGACCAAGTGATGAAAAAACTGAA
351 TGCAATTGCTGAGGAACGAAAGAATTTTCATTTGCAAGAAAAGATTATAG
401 AGAGACAAGCTGCTACACGGGAAACAG**GTACTCATCTTAAATTAGTATTA**
451 **CAACTTAGTTTATATTCATTTGTTTTGGGCAATGATCAAATTATGTAAAG**
501 **GTCAAATATACTCATGTACTACTGAAAATAGTTTAAATATACCTCTAGTT**
551 **ATACTATTAGTACGAACATACTCCTCCCATATACTTTGGAACAAATATTC**
601 **CCTTAACGAAATAAGACACGTGAAAAGTTCAGATTCAAATTATCCACCCT**
651 **CAATTTTAAGATCTGATTTCTTTAGGAAACCACTCATCTCCTCCGTTTTG**
701 **AGTTCTTAACGAAGCAGCTCAGAGAAAAGAGGTTTTCTTCTGTTCTGTTT**
751 **CTGCTGCATTTGTGTCTTAATCCAATAACAAACAATACAAATTAATATTA**
801 **TGTTCACGATGAGGGTAGTCTTTCTAGCTAGACATGAACTGAGTGTAAAT**
851 **TTTGTTTTAAGGAAGAAAAAGAAATGATTAGGCTGGATTTCTTTCAGAGT**
901 **GGAATATAGGGGGATAAAGTTGGAGCATAGAGTTCCATCGTTTATTTCTT**
951 **TCCTTAAAGTAACAAGTTCAACAAAATGATATCAAGGTACGGTAATGGAA**
1001 **AATTATTAGACACGTCTAAACTACAAAAATGGAATAGAAACTTAAATTAT**
1051 **CAGTGACAATATCATCCTTTAATAAAGCTACCAAATTTAAATCATGATAC**
1101 **AGAGAAGAAACCAAAAAAATTAGGGGTGAATTATTTGATTCTATGCTTAT**
1151 **CACATGTCTTCCCATCAACATCAAAGGAAAAATTGTGCCAAAGTATAAAC**
1201 **GGTGCGGTATATTTGGATTGAAAGTAAAACAGGAGGATACATTTGGACTA**
1251 **AAAGTATAACAATAAGTATATTTGATCATTTTATGTATCAAATTCATGTG**
1301 **GTTTTTGGGGAGAAGGGAAGTTTCAATGTTTTCAATCTGCTCCTCATCTC**
1351 **ATCCATATCTCTTTATTGTGCAAAACCCTTCTCTATTTAACTATTTTCTG**
1401 **CCGACTCCTAATGAGCTTGAATGTAACAATATTCTCATCTGGACATTGCT**

1451 **TGCACCAG**GTTCTGTGTTAACTGAACCACAAGTTTATGGAAGGGACAAAG
1501 AAAAAGATGAGATAGTGAAAATCCTAATAAACAATGTTAGTGATGCCCAA
1551 AAACTCTCAGTCCTCCCAATACTTGGTATGGGGGGACTAGGAAAGACAAC
1601 TCTTTCCCAAATGGTCTTCAATGATCAGAGAGTAACTGAGCGTTTCTATC
1651 CCAAAATATGGATTTGCGTCTCGGATGATTTTGATGAGAAGAGGTTGATA
1701 AAGGCAATAGTAGAATCTATTGAAGGGAAGTCCCTCAGTGACATGGACTT
1751 GGCTCCACTTCAAAAGAAGCTTCAAGAGTTGCTGAATGGAAAAAGATACT
1801 TCCTTGTCTTAGATGATGTTTGGAATGAAGATCAACATAAGTGGGCTAAT
1851 TTAAGAGCAGTCTTGAAGGTTGGAGCAAGTGGTGCATTTGTTCTAACTAC
1901 TACTCGTCTTGAAAAGGTTGGATCAATTATGGGAACATTGCAACCATATG
1951 AATTGTCAAATCTGTCTCCAGAGGATTGTTGGTTTTTGTTCATGCAGCGT
2001 GCATTTGGACACCAAGAAGAAATAAATCCAAACCTTGTGGCAATCGGAAA
2051 GGAGATTGTGAAAAAATGTGGTGGTGTGCCTCTAGCAGCCAAGACTCTTG
2101 GAGGTATTTTGCGCTTCAAGAGAGAAGAAAGAGAATGGGAACATGTGAGA
2151 GACAGTCCGATTTGGAATTTGCCTCAAGATGAAAGTTCTATTCTGCCTGC
2201 CCTGAGGCTTAGTTACCATCATCTTCCACTTGATTTGAGACAATGCTTTG
2251 TGTATTGTGCGGTATTCCCAAAGGACACCAAAATGGCAAAGGAAAATCTT
2301 ATCGCTTTTTGGATGGCACATGGTTTTCTTTTATCGAAAGGAAATTTGGA
2351 GCTAGAGGATGTAGGTAATGAAGTATGGAATGAATTATACTTGAGGTCTT
2401 TCTTCCAAGAGATTGAAGTTGAATCTGGTAAAACTTATTTCAAGATGCAT
2451 GACCTCATCCATGATTTGGCTACATCTCTGTTTTCAGCAAACACATCAAG
2501 CAGCAATATTCGTGAAATAAATGCTAATTATGATGGATATATGATGTCGA
2551 TTGGTTTTGCTGAAGTGGTATCTTCTTACTCTCCTTCACTCTTGCAAAAG
2601 TTTGTCTCATTAAGGGTGCTTAATCTAAGAAACTCGAACCTAAATCAATT
2651 ACCATCTTCCATTGGAGATCTAGTACATTTAAGATACCTGGACTTGTCTG
2701 GCAATTTTAGAATTCGTAATCTTCCAAAGAGATTATGCAGGCTTCAAAAT
2751 CTGCAGACTCTTGATCTACATTATTGCGACTCTCTTTCTTGTTTGCCAAA
2801 ACAAACAAGTAAACTTGGTAGTCTCCGAAATCTTTTACTTGATGGCTGTT
2851 CATTGACGTCAACGCCACCAAGGATAGGATTGTTGACATGCCTTAAGTCT
2901 CTAAGTTGCTTTGTTATTGGCAAGAGAAAAGGTTATCAACTTGGTGAACT
2951 AAAAAACCTAAATCTCTATGGCTCAATTTCAATCACAAAACTTGACAGAG
3001 TGAAGAAAGATAGCGATGCAAAAGAAGCTAATTTATCTGCTAAAGCAAAT

3051 CTGCACTCTTTATGCCTGAGTTGGGACCTTGATGGAAAACATAGATATGA
3101 TTCAGAAGTTCTTGAAGCCCTCAAACCACACTCCAATCTGAAATATTTAG
3151 AAATCAATGGCTTCGGAGGAATCCGTCTCCCAGATTGGATGAATCAATCA
3201 GTTTTGAAAAATGTTGTCTCTATTAGAATTAGAGGTTGTGAAAACTGCTC
3251 ATGCTTACCACCCTTTGGTGAGCTGCCTTGTCTAGAAAGTCTAGAGTTAC
3301 ACACCGGGTCAGCAGATGTGGAGTATGTTGAAGATAATGTTCATCCTGGA
3351 AGGTTTCCATCCTTGAGGAAACTTGTTATATGGGACTTTAGTAATCTAAA
3401 AGGATTGCTGAAAAAGGAAGGAGAAAAGCAATTCCCTGTGCTTGAAGAGA
3451 TGACATTTTACTGGTGCCCTATGTTTGTTATTCCGACCCTTTCTTCTGTC
3501 AAGACATTGAAAGTTATTGCGACAGATGCAACAGTTTTGAGGTCCATATC
3551 TAATCTTAGGGCTCTTACTTCCCTTGACATTAGCAATAACGTAGAAGCTA
3601 CTTCACTCCCAGAAGAGATGTTCAAAAGCCTTGCAAATCTCAAATACTTG
3651 AATATCTCTTTCTTTAGGAATCTCAAAGAGTTGCCTACCAGCCTGGCTAG
3701 TCTCAATGCTTTGAAGAGTCTCAAATTTGAATTTTGTAACGCACTAGAGA
3751 GTCTCCCAGAGGAAGGGGTGAAAGGTTTAACTTCACTCACCGAGTTGTCT
3801 GTCAGTAACTGTATGATGCTAAAATGTTTACCGGAGGGATTGCAGCACCT
3851 AACAGCCCTCACAACTTTAACAATTACTCAATGTCCAATAGTATTCAAGC
3901 GGTGTGAGAGAGGAATAGGAGAAGACTGGCACAAAATTGCTCACATTCCA
3951 TATTTGACTCTATATGAGTGA

Figure 6F

1 ATGGCGGAAGCTTTTCTTCAAGTTCTGCTAGAAAATCTCACTTCTTTCAT
51 CGGAGATAAACTTGTATTGATTTTCGGTTTCGAAAAGGAATGTGAAAAGC
101 TGTCGAGTGTGTTTTCCACAATTCAAGCTGTGCTTCAAGATGCTCAGGAG
151 AAGCAATTGAAGGACAAGGCAATTGAGAATTGGTTGCAGAAACTCAATTC
201 TGCTGCCTATGAAGTTGATGATATATTGGGCGAATGTAAAAATGAGGCAA
251 TAAGATTTGAGCAGTCTCGATTAGGGTTTTATCACCCAGGGATTATCAAT
301 TTCCGTCACAAAATTGGGAGAAGGATGAAAGAGATAATGGAGAAACTAGA
351 TGCAATATCTGAGGAAAGAAGGAAGTTTCATTTCCTTGAAAAAATTACAG
401 AGAGACAAGCTGCCGCTGCTACGCGTGAAACAG**GTGTGAGTACTGAGTAA**
451 **TTGTAGCTTAGTTAATATTCAATTTGTTACCACATCATGTGTTCACCGTG**
501 **ATCTCTACAG**TAGGATGGCAATGGGGCTGGGCGAGGTTGGAG**GTGTGCAG**
551 **GTGTGTGGCGCAACCCCAACTTTGAGTCTACATAAGTAGGTACTTAAATT**
601 **TGTATAGAGTTGAACAAG**TACAAACGCCTCCTACTTGGTGTCCTTATGCG
651 TATTATGTCACTTAGGATGCATGTGTCTACTTGTTCAACTTTATATGAGT
701 TTAAGTTCTACTTGTGCACACCCAAAGTTGGAGCGCGTAGAT**GTCAGTTG**
751 **ATACCAAGTTAAAAAGGCATATTTATGAATTATGCCTTTAAATTATGATT**
801 **CAATTTTGTATCAGTCTGTCCAAAATATGTTCTAGTGAAAGTGTTAAACT**
851 **TAGTCTGGATCTGCTATTGAAAGTGAATTTTTGTGGCACTAAACAATGCA**
901 **ATGGGTCTGGATTCATTTTTGCATTAACTTTTGTTTAGACGATTTTCTTT**
951 **ATCGAATTTTACTGTCTAAAATGGAAAAAGCAAAGAAATAAGAAGTATAC**
1001 **AGAGGCTGACTTCTTCATAGTATCTATCATATAAAAAAAAGCATTGATTA**
1051 **CTAGGATATGGGTTCTTTTAAATTACAAATTTGTGAGTTAAAACAGTTCT**
1101 **GTTGGGAAGGATTTAGATACACGTGGATAGTATCTAGAAGTTTTTTAAAT**
1151 **AAAAAATTAGCAAATTATGCGGGCTGGGGCGGGTTGAAAACAGCAAACTT**
1201 **TGCAAGGCTTGGCGGGTCGAAATCTTTGCAAGTTTGTGTGGGTTTGCCCT**
1251 **GCACCACCCAATCTGCCATTCCTGTCTAAATGTTTGTTTTGTCTATAATT**
1301 **CTTGCTGACTCATTCTAATGAGCTCAATTGTAACAAATTCTTTGTGTCCA**
1351 **CACTACTTGGAACAG**GTTTTGTGTTAACTGAACCAAAAGTCTACGGAAGG
1401 GACAAAGAGGAGGATGAGATAGTGAAAATTCTGATAAACAATGTTAATGT

1451 TGCCGAAGAACTTCCAGTCTTCCCTATAATTGGTATGGGGGGACTAGGAA
1501 AGACGACACTTGCCCAAATGATCTTCAACGATGAGAGAGTAACTAAGCAT
1551 TTCAATCCCAAAATATGGGTTTGTGTCTCAGATGATTTTGATGAGAAGAG
1601 GTTAATTAAGACAATTATAGGAAATATTGAAAGAAGTTCTCCTCATGTTG
1651 AGGACTTGGCTTCATTTCAGAAGAAGCTCCAGGAGTTATTGAATGGAAAA
1701 CGATACTTGCTTGTCTTAGATGATGTTTGGAATGATGATCTAGAAAAGTG
1751 GGCTAAGTTAAGAGCAGTCTTAACTGTTGGAGCAAGAGGTGCTTCTATTC
1801 TAGCTACTACTCGTCTTGAAAAGGTTGGATCAATTATGGGAACGTTGCAA
1851 CCATATCATTTGTCAAATTTGTCTCCACATGATAGTTTACTTTTGTTTAT
1901 GCAACGCGCATTTGGGCAACAAAAAGAAGCAAATCCTAATCTAGTGGCCA
1951 TTGGAAAGGAGATTGTGAAGAAATGTGGTGGTGTGCCTTTAGCAGCCAAG
2001 ACTCTTGGTGGTCTTTTACGCTTCAAGAGAGAAGAGAGTGAATGGGAACA
2051 TGTGAGAGATAATGAGATTTGGAGTCTGCCTCAAGATGAAAGTTCTATTT
2101 TGCCTGCTCTAAGACTGAGTTATCATCACCTTCCACTTGATTTGAGACAA
2151 TGCTTTGCGTATTGTGCAGTATTCCCAAAGGACACCAAAATGATAAAGGA
2201 AAATCTCATTACTCTCTGGATGGCGCATGGTTTTCTTTTATCAAAGGGAA
2251 ACTTGGAGCTAGAGGATGTGGGTAATGAAGTATGGAATGAATTATACTTG
2301 AGGTCTTTCTTCCAAGAAATTGAAGCTAAATCGGGTAATACTTATTTCAA
2351 GATACATGATCTAATCCATGATTTGGCTACATCTCTGTTTTCGGCAAGCG
2401 CATCATGCGGCAATATCCGCGAAATAAATGTCAAAGATTATAAGCATACA
2451 GTGTCCATTGGTTTCGCTGCAGTGGTGTCTTCTTACTCTCCTTCGCTCTT
2501 GAAAAAGTTTGTCTCGTTAAGGGTGCTTAATCTAAGTTACTCAAAACTTG
2551 AGCAATTACCGTCTTCCATTGGAGATCTATTACATTTAAGATACCTGGAC
2601 CTGTCTTGCAATAACTTCCGTAGTCTTCCAGAGAGGTTGTGCAAGCTTCA
2651 AAATCTTCAGACTCTTGATGTACATAATTGCTACTCACTTAATTGTTTGC
2701 CAAAACAAACAAGTAAACTTAGTAGTCTCCGACATCTTGTTGTTGATGGC
2751 TGTCCATTGACTTCTACTCCACCAAGGATAGGATTGTTGACATGCCTTAA
2801 GACTCTAGGTTTCTTTATTGTGGGAAGCAAGAAAGGTTATCAACTTGGTG
2851 AACTGAAAAACCTAAATCTCTGCGGCTCAATTTCAATCACACACCTTGAG
2901 AGAGTGAAGAACGATACGGATGCAGAAGCCAATTTATCTGCAAAAGCAAA
2951 TCTGCAATCTTTAAGCATGAGTTGGGATAACGATGGACCAAACAGATATG
3001 AATCCAAAGAAGTTAAAGTGCTTGAAGCACTCAAACCACACCCCAATCTG

3051 AAATATTTAGAGATCATTGCCTTCGGAGGATTCCGTTTTCCAAGCTGGAT
3101 AAATCACTCAGTTTTGGAGAAGGTCATCTCTGTTAGAATTAAAAGCTGCA
3151 AAAACTGCTTGTGCTTACCACCCTTTGGGGAGCTTCCTTGTCTAGAAAAT
3201 CTAGAGTTACAAAACGGATCTGCGGAGGTGGAGTATGTTGAAGAGGATGA
3251 TGTCCATTCTAGATTCTCCACAAGAAGAAGCTTTCCATCCCTGAAAAAAC
3301 TTCGTATATGGTTCTTTCGCAGTTTGAAAGGGCTGATGAAAGAGGAAGGA
3351 GAAGAGAAATTCCCCATGCTTGAAGAGATGGCGATTTTATATTGCCCTCT
3401 GTTTGTTTTTCCAACCCTTTCTTCTGTCAAGAAATTAGAAGTTCACGGCA
3451 ACACAAACACTAGAGGTTTGAGCTCCATATCTAATCTTAGCACTCTTACT
3501 TCCCTCCGCATTGGTGCTAACTACAGAGCGACTTCACTCCCAGAAGAGAT
3551 GTTCACAAGTCTTACAAATCTCGAATTCTTGAGTTTCTTTGACTTCAAGA
3601 ATCTCAAAGATCTGCCTACCAGCCTGACTAGTCTCAATGCTTTGAAGCGT
3651 CTCCAAATTGAAAGTTGTGACTCACTAGAGAGTTTCCCTGAACAAGGGCT
3701 AGAAGGTTTAACTTCACTCACACAGTTGTTTGTTAAATACTGTAAGATGC
3751 TAAAATGTTTACCCGAGGGATTGCAGCACCTAACAGCCCTCACAAATTTA
3801 GGAGTTTCTGGTTGTCCAGAAGTGGAAAAGCGCTGTGATAAGGAAATAGG
3851 AGAAGACTGGCACAAAATTGCTCACATTCCAAATCTGGATATTCATTAG

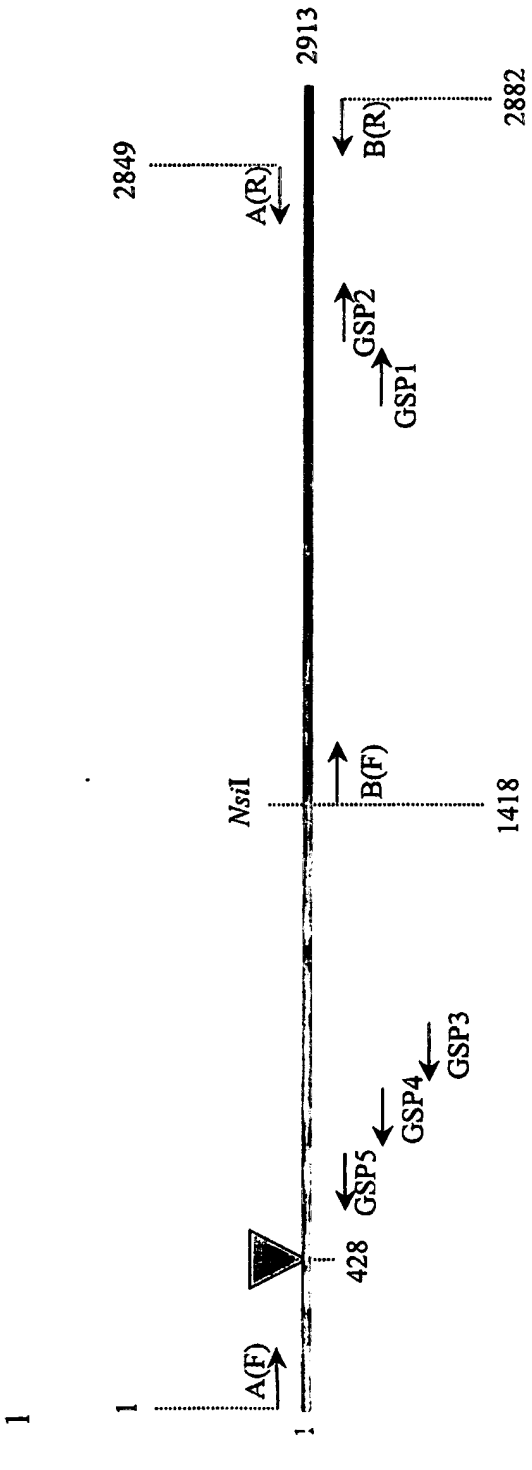
1
1
2913
2882
2849
A(R)
B(R)
GSP2
GSP1
NsiI
B(F)
1418
GSP3
GSP4
GSP5
428
A(F)
intron
primer

Figure 7

A
MAEAFIQVLLDNLTSFLKGELVLLFGFQDEFQRLSSMFSTIQAVLEDAQEKQLNN 55
KPLENWLQKLNAATYEVDDILDEYKTKATRFSQSEYGRYHPKVIPFRHKVGKRMD 110
QVMKKLKAIAEERKNFHLHEKIVERQAVRRETGSVLTEPQVYGRDKEKDEIVKIL 165

B
INNVSDAQHLSVL*pilgmgglgkttla*QMVFNDQRVTEHFHSKIWICVSEDFDEK 220
RLIKAIVESIEGRPLLGEMDLAPLQKKLQELLNG*kryllvlddvw*NEDQQKWANL 275
RAVLKVGASGA*svltttr*LEKVGSIMGTLQPYELSNLSQEDCWLLFMQRAFGHQE 330
EINPNLVAIGKEIVKKSGGVPLAAKTLGGILCFKREERAWEHVRDSPIWNLPQDE 385
SSILPALRLSYHQLPLDLKQCFAYCAVFPKDAKMEKEKLISLWMAHGFLLSKGNM 440
ELEDVGDEVWKELYLRSFFQEIEVKDGKTYFK*mhdlihdlat*SLFSANTSSSNIR 495
EINKHS 501

C
YTHMMSIGFAEVVFFYTLPPLEK 524
FISLRVLNLGDST.FNKLPSSIGD 547
LVHLRYLNLYGSG.MRSLPKQLCK 570
LQNLQTLDLQYCTKLCCLPKETSK 594
LGSLRNLLLDGSQSLTCMPPRIGS 618
LTCLKTLGQFVVGRKKGYQ 637
LGELGNLNLYGSIKISHLERVKNDKDAKEANLSA 671
KGNLHSLSMSWNNFGPHIYESEEVKVLEALKP 703
HSNLTSLKIYGFRGIH.LPEWMNHSV 728
LKNIVSILISNFRNCSCLPPFGD 751
LPCLESLELHWGSAD 766
VEYVEEVDIDVHSGFPTRIR 786
FPSLRKLDIWDFGSLKGLLKKEGEEQ 812
FPVLEEMIIHECPFLT.LSSN 832
LRALTSLRICYNKVATSFPEEMFKN 857
LANLKYLTISRCNNLKELPTSLAS 881
LNALKSLKIQLCCALESLPEEGLEG 906
LSSLTELFVEHCNMLKCLPEGLQH 930
LTTLTSLKIRGCPQLIKRCEKGIGEDWHK 959
ISHIPNVNIYI 970

L..L..L.L..C..α..αP.. LRR consensus
N
S

Figure 8

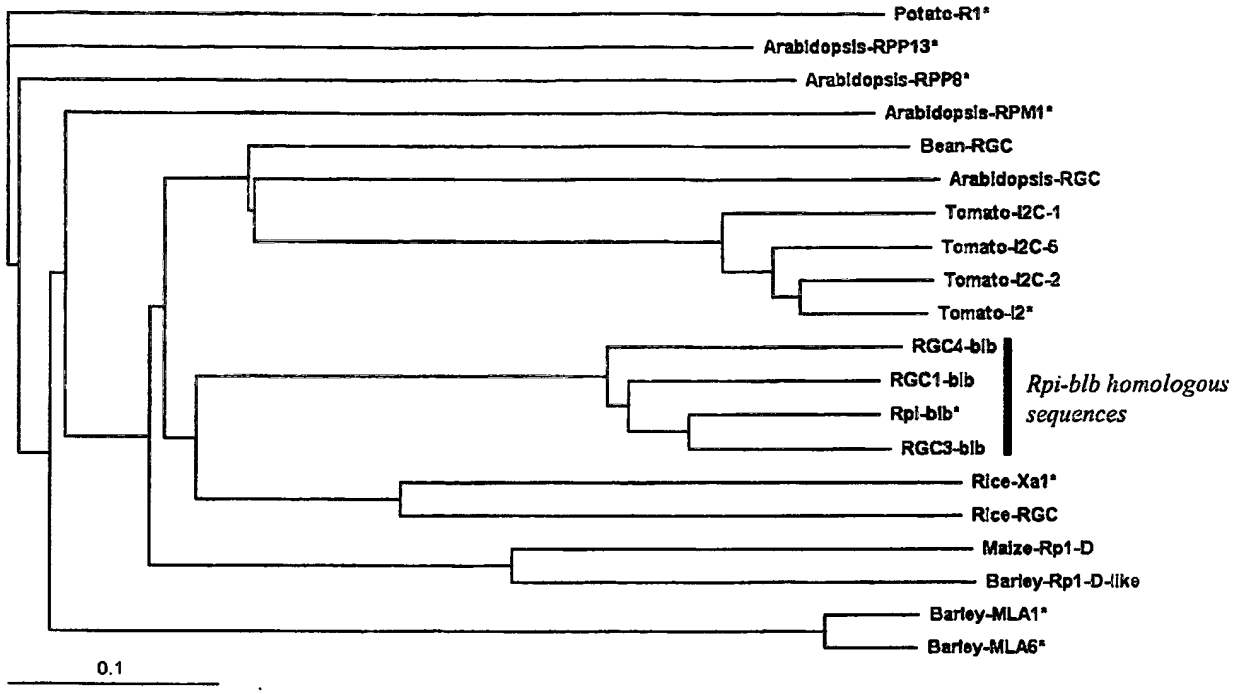
Potato-R1*
Arabidopsis-RPP13*
Arabidopsis-RPP8*
Arabidopsis-RPM1*
Bean-RGC
Arabidopsis-RGC
Tomato-I2C-1
Tomato-I2C-5
Tomato-I2C-2
Tomato-I2*
RGC4-blb
RGC1-blb
Rpi-blb*
RGC3-blb
Rpi-blb homologous sequences
Rice-Xa1*
Rice-RGC
Maize-Rp1-D
Barley-Rp1-D-like
Barley-MLA1*
Barley-MLA6*
0.1

Figure 9A

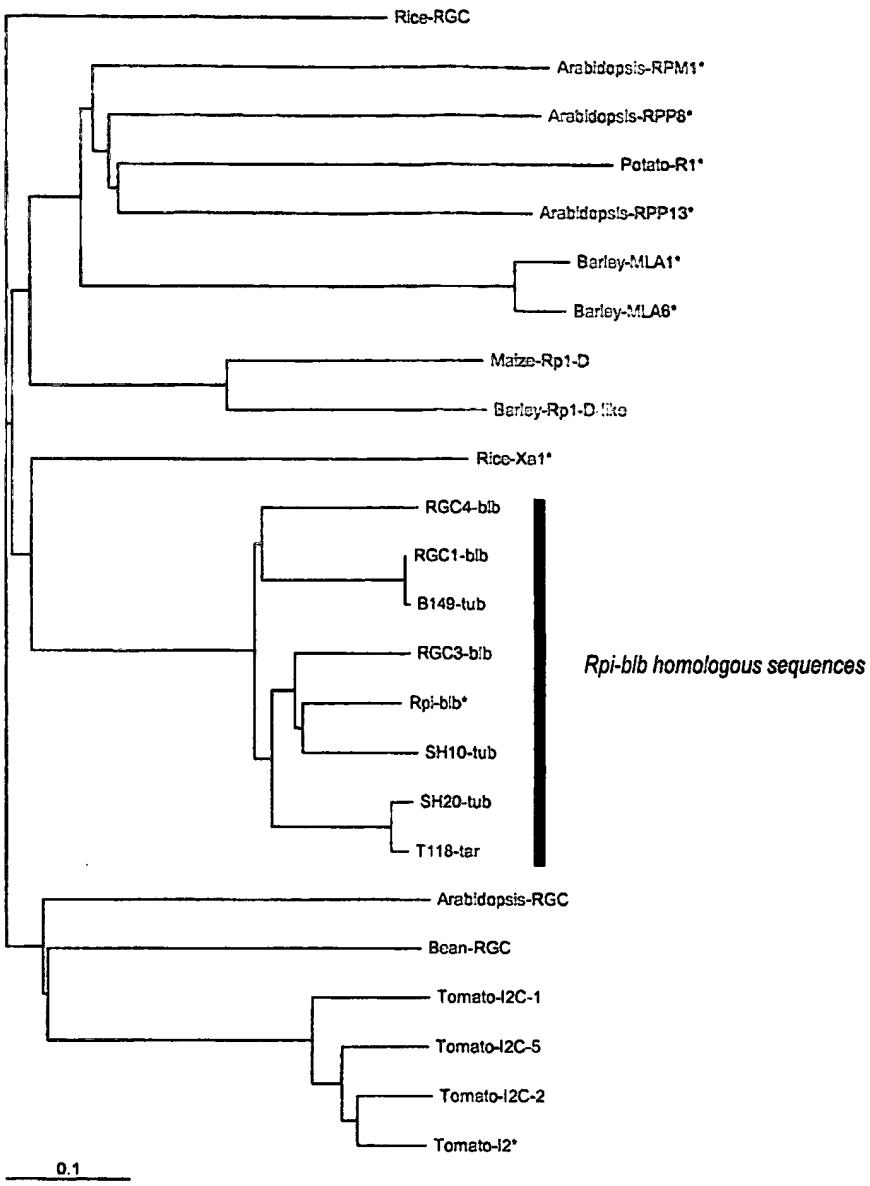
Rice-RGC
Arabidopsis-RPM1*
Arabidopsis-RPP8*
Potato-R1*
Arabidopsis-RPP13*
Barley-MLA1*
Barley-MLA6*
Maize-Rp1-D
Barley-Rp1-D like
Rice-Xa1*
RGC4-blb
RGC1-blb
B149-tub
RGC3-blb
Rpi-blb*
SH10-tub
SH20-tub
T118-tar
Rpi-blb homologous sequences
Arabidopsis-RGC
Bean-RGC
Tomato-I2C-1
Tomato-I2C-5
Tomato-I2C-2
Tomato-I2*
0.1

Figure 9B

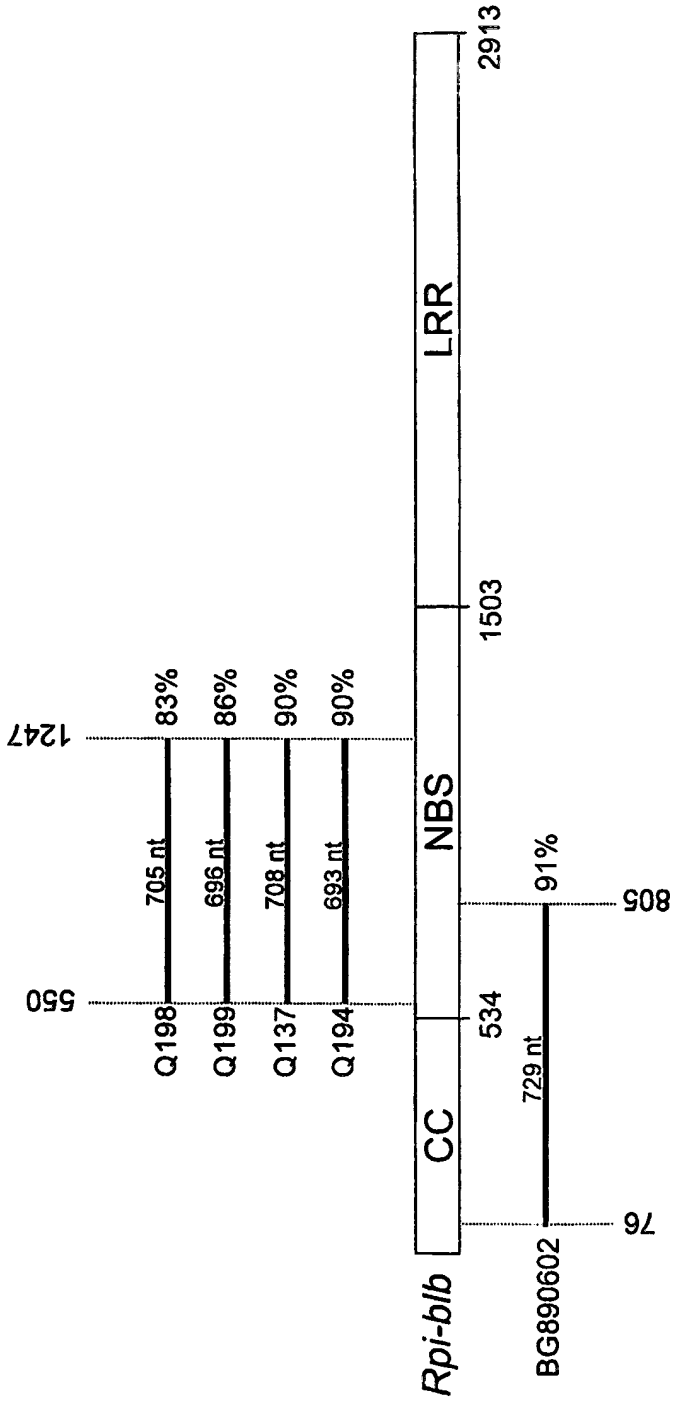
Rpi-blb
CC
NBS
LRR
2913
1503
534
76
805
550
1247
BG890602
729 nt
91%
Q198
Q199
Q137
Q194
705 nt
696 nt
708 nt
693 nt
83%
86%
90%
90%

Figure 9C

```
Rpi-blb     MAEAFIQVLLDNLTSFLKGELVLLFGFQDEFQRLSSMFSTIQAVLEDAQEKQLNNKPLEN   60
RGC3-blb          V                                              D         60
RGC1-blb       L       F IQ   G V   EK  KK      M                KY AIK   60
RGC4-blb       L    E     IGDK   I   EK CEK   V        Q        KD AI     60

Rpi-blb     WLQKLNAATYEVDDILDEYKTKATRFSQSEYGRYHPKVIPFRHKVGKRMDQVMKKLKAIA   120
RGC3-blb                                L                              N   120
RGC1-blb       V A        DC  E A  K AVL     RT T CY       KEM E  D       120
RGC4-blb       S A       G C NE I  E  RL F   GI N    I R  KEI E  D  S     120

Rpi-blb     EERKNFHLHEKIVERQAVR--RETG----------------------------------   143
RGC3-blb          Q   I     AT--     ----------------------------------   143
RGC1-blb    R     D R I     A -- Q  ----------------------------------   143
RGC4-blb    RK  FL   T     AAAT   VGWQWGWARLEYKRLLLGVLMRIMSLRMHVSTCSTL   180

Rpi-blb     -----------------SVLTEPQVYGRDKEKDEIVKILINNVSDAQHLSVLpilgmggl   186
RGC3-blb    -----------------                              K              186
RGC1-blb    -----------------F     K    E  E             YSEEVP         186
RGC4-blb    YEFKFYLCTPKVGARRCF     K       E            NV EE P F  I     240


Rpi-blb     gkttlaQMVFNDQRVTEHFHSKIWICVSEDFDEKRLIKAIVESIEGRPLLGEMDLAPLQK   246
RGC3-blb         S          R YP       D                 KS S-D          245
RGC1-blb                 I    NL   V   D                 KS G-D          245
RGC4-blb           I   E   K  NP   V   D         T IGN  - SSPHVE   SF     299

Rpi-blb     KLQELLNGkryllvlddvwNEDQQKWANLRAVLKVGASGAsvltttrLEKVGSIMGTLQP   306
RGC3-blb              F           H                 F                     305
RGC1-blb              F           E  D       I      I I      I        L   305
RGC4-blb                        D LE   K     T   R   I A                   359

Rpi-blb     YELSNLSQEDCWLLFMQRAFGHQEEINPNLVAIGKEIVKKSGgvplaaktlggILCFKRE   366
RGC3-blb          P    F                              C              R      365
RGC1-blb    Q             K    C  T TS K ME         C            L R      365
RGC4-blb    H     PH SL         Q K A                C            L R      419

Rpi-blb     ERAWEHVRDSPIWNLPQDESSILPALRLSYHQLPLDLKQCFAYCAVFPKDAKMEKEKLIS   426
RGC3-blb      E                              H     R   V        T  A  N  A   425
RGC1-blb    SE       E        N V           H     R            T I   Y  A   425
RGC4-blb    SE      NE  S                   H     R            T  I  N  T   479

Rpi-blb     LWMAHGFLLSKGNMELEDVGDEVWKELYLRSFFQEIEVKDGKTYFKmhdlihdlatSLFS   486
RGC3-blb    F            L      N   N             ES                       485
RGC1-blb         S              N   N              S                   M   485
RGC4-blb                 L      N   N            A S N    I                539

Rpi-blb     ANTSSSNIREINKH----------SYTHMMSIGFAEVVFFYTLPPLEKFISLRVLNLGDS   536
RGC3-blb                 AN----------YDGY          SS SPSL Q  V        RN   535
RGC1-blb     SA  RS  Q  VKDDEDMMFIVTN KD      S   SS SPSLFKR V        SN   545
RGC4-blb     SA CG      VK----------D K TV     A  SS SPSL K  V        SY   589

Rpi-blb     TFNKLPSSIGDLVHLRYLNLYG-SGMRSLPKQLCKLQNLQTLDLQYCTKLCCLPKETSKL   595
RGC3-blb    NL Q              D S NFRI N   R            H  DS S    Q       595
RGC1-blb    E EQ     V        D S -NKIC    R            YN QS S    Q       604
RGC4-blb    KLEQ        L     D SC-NNF    ER           VHN YS N    Q       648

Rpi-blb     GSLRNLLLDGSQSLTCMPPRIGSLTCLKTLGQFVVGRKKGYQLGELGNLNLYGSIKISHL   655
RGC3-blb              -C   ST     L     S SC  I KR        K         S TK   654
RGC1-blb    C     V  H-CP  S      L        Y    ER        R    R A S T    663
RGC4-blb    S   H VV  -CP  ST     L        F I  S         K    C   S T    707

Rpi-blb     ERVKNDKDAKEANLSAKGNLHSLSMSWNNFGPHIYESEEVKVLEALKPHSNLTSLKIYGF   715
RGC3-blb    D   K S           A     CL  DLD K R D ---E          KY E N     711
RGC1-blb         ME           A         D--R NR               P  KY E ID   721
RGC4-blb         T  -         A  Q      D D  NR   K            P  KY E IA   766
```

```
Rpi-blb     RGIHLPEWMNHSVLKNIVSILISNFRNCSCLPPFGDLPCLESLELHWGSADVEYVEEVDI      775
RGC3-blb    G  R  D   Q     V   R RGCE         E          T         DN--     769
RGC1-blb    C FC  D         V      GCE         E         QD  VE     DS--     779
RGC4-blb    G FRF S I     EKVI VR KSCK  L      E     N   QN   E      D--     824

Rpi-blb     DVHSGFPTRIRFPSLRKLDIWDFGSLKGLLKKEGEEQFPVLEEMIIHECPFLTLS-----     830
RGC3-blb    -  P -----         V    SN          K        TFYW  MFVIPTLSSV     823
RGC1-blb    ----  L  R         H GG CN    QRMK A         K SD  MFVFPTLSSV     835
RGC4-blb        R S  RS    K  R  F R     M E    K  M    A LY  LFVFPTLSSV     884

Rpi-blb     -----------------SNLRALTSLRICYNKVATSFPEEMFKNLANLKYLTISRCNNLK     873
RGC3-blb    KTLKVI-ATDATVLRSI         D SN VE   L      S       N  FFR       882
RGC1-blb    KKLEIWGEADAGGLSSI   ST    K FS HTV  LL       E  I  SV FLE       895
RGC4-blb    KKLEVHGNTNTRGLSSI   ST      GA YR   L     TS T  EF SFFDFK       944

Rpi-blb     ELPTSLASLNALKSLKIQLCCALESLPEEGLEGLSSLTELFVEHCNMLKCLPEGLQHLTT     933
RGC3-blb                  FEF N         VK  T      S SN M              A     942
RGC1-blb            N  C D RY Y                                               955
RGC4-blb    D     T      R Q ES DS   F  Q     T   Q   KY K              A    1004

Rpi-blb     LTSLKIRGCPQLIKRCEKGIGEDWHKISHIPNVNIYI                          970
RGC3-blb      T T TQ  IVF    R         A   YLTL E                           979
RGC1-blb                                                                    992
RGC4-blb      N GVS   EVE   D E        A    LD H-                          1040
```

Figure 10A

```
Rpi-blb    MAEAFIQVLLDNLTSFLKGELVLLFGFQDEFQRLSSMFSTIQAVLEDAQEKQLNNKPLEN   60
RGA3-blb          V                                                D      60
SH10-tub           I                    N       I                   D      60
RGA1-blb       L        F IQ   G V   EK  KK      M                KY AIK   60
B149-blb                F IQ   G V   EK  KK      M                KY AIK   60
SH20-tub             E I   IQ   G  L  END ENI  R                   KD AIK   60
T118-tar             E I   IQ   G  L  EN  ENI  R                   KD AIK   60
RGA4-blb       L     E     IGDK   I   EK CEK   V        Q          KD AIE   60

Rpi-blb    WLQKLNAATYEVDDILDEYKTKATRFSQSEYGRYHPKVIPFRHKVGKRMDQVMKKLKAIA  120
RGA3-blb                                L                             N   120
SH10-tub                                   AY                         N   120
RGA1-blb         V A         DC  E A  K AVL     RT T CY       KEM E  D   120
B149-blb         V A         DC  E A  K AVL     RT T CY       KEM E  D   120
SH20-tub           V K   L    C -- A LE  RL CH    A V    I    IKEM E  D   118
T118-tar           A K   L    C -- A LE  RL  H    A V    I    IKEM E  D   118
RGA4-blb         S A       G C NE I  E  RL F   GI N    I R   KEI E  D  S  120

Rpi-blb    EERKNFHLHEKIVERQAVR--RETG----------------------------------   143
RGA3-blb           Q   I     AT--     ----------------------------------   143
SH10-tub               I       --     ----------------------------------   143
RGA1-blb      R    D R I     A -- Q   ----------------------------------   143
B149-blb      R    D R I     A -- Q   ----------------------------------   143
SH20-tub   K  TD       I    VA --P   ----------------------------------   141
T118-tar   K  TD       I    VA --P   ----------------------------------   141
RGA4-blb      RK  FL   T     AAAT    VGWQWGWARLEYKRLLLGVLMRIMSLRMHVSTCSTL  180

Rpi-blb    -----------------SVLTEPQVYGRDKEKDEIVKILINNVSDAQHLSVLpilgmggl  186
RGA3-blb   -----------------                                 K          186
SH10-tub   -----------------                 E                          186
RGA1-blb   -----------------F     K    E  E            YSEEVP           186
B149-blb   -----------------F     K    E  E            YSEEVP           186
SH20-tub   -----------------F          E               N  E             184
T118-tar   -----------------P          E               N LE             184
RGA4-blb   YEFKFYLCTPKVGARRCF     K       E           NV EE P F  I       240

Rpi-blb    gkttlaQMVFNDQRVTEHFHSKIWICVSEDFDEKRLIKAIVESIEGRPLLGEMDLAPLQK  246
RGA3-blb       S           R YP       D                KS S-D          245
SH10-tub               I                        L   I                   246
RGA1-blb               I    NL   V   D                KS G-D           245
B149-blb               I    NL   V   D                KS G-D           245
SH20-tub                    YP       D        EN IGN  - SS DVK   SF     243
T118-tar                    YP       D        ET IGN  - SS DVK   SF     243
RGA4-blb         I   E   K  NP   V   D         T IGN  - SSPHVE   SF     299

Rpi-blb    KLQELLNGkryllvlddvwNEDQQKWANLRAVLKVGASGAsvltttrLEKVGSIMGTLQP  306
RGA3-blb             F          H               F                        305
SH10-tub             F                          F  A                     306
RGA1-blb             F          E  D       I     I I      I        L    305
B149-blb             F          E  D       I     I I      I        L    305
SH20-tub   Q                       D   V                                 303
T118-tub   Q                       D   A                                 303
RGA4-blb                   D LE   K     T   R   I A                      359

Rpi-blb    YELSNLSQEDCWLLFMQRAFGHQEEINPNLVAIGKEIVKKSGgvplaaktlggiLCFKRE  366
RGA3-blb         P    F                              C              R     365
SH10-tub                I C                                        R     366
RGA1-blb   Q             K    C  T TS K ME        C             L R     365
B149-blb   Q             K    C  T TS K ME        C             L R     365
SH20-tub   Q       D     I    R     S                            L R     363
T118-tar   Q       D     I   YR     S                            L R     363
RGA4-blb   H     PH SL         Q K A              C             L R     419

Rpi-blb    ERAWEHVRDSPIWNLPQDESSILPALRLSYHQLPLDLkqcfaycavfPKDAKMEKEKLIS  426
RGA3-blb                                  H     R   V        T  A  N  A   425
SH10-tub             E      E R           H     R            T           426
RGA1-blb   SE        E        N V         H     R            T I   Y  A   425
B149-blb   SE        E        N V         H     R            T I   Y  A   425
SH20-tub   K E       E        M           H   A R            T    K V     423
T118-tar   K E       E        M    V      H     R            T    K V     423
RGA4-blb    SE      NE  S                 H     R            T  I  N  T   479

Rpi-blb    LWMAHGFLLSKGNMELEDVGDEVWKELYLRSFFQEIEVKDGNTYFKmhdlihdlATSLFS  486
RGA3-blb   F           L          N              ES                       485
SH10-tub             LE KLQP     N  S   C         A C          H          486
RGA1-blb        S               N   N              S                 M    485
```

```
B149-blb        S              N   N         G    S                  M     485
SH20-tub             RR L      RN G N              RY N                    483
T118-tar             RR L       N   N              RY N                    483
RGA4-blb               L        N   N            A S N    I               539

Rpi-blb    ANTSSSNIREINKH----------SYTHMMSIGFAEVVFFYTLPPLEKFISLRVLNLGDS   536
RGA3-blb                AN----------YDGY          SS SPSL Q  V       RN   535
SH10-tub   S            VK----------G P K     T   SS SPSLSQ  V       SNL  537
RGA1-blb   SA  RS  Q    VKDDEDMMFIVTN KD      S   SS SPSLFKR V       SN   545
B149-blb   SA  RS  Q    VKDDEDMMFIVTN KD      S   SS SPSLFKR V       SN   545
SH20-tub                VE----------          S   SS SPSL Q  V       SY   534
T118-tar                VE----------          S   SS SPSL Q  V       SY   534
RGA4-blb   SA CG        VK----------D K TV     A  SS SPSL K  V       SY   589

Rpi-blb    TFNKLPSSIGDLVHLRYLNLYG-SGMRSLPKQLCKLQNLQTLDLQYCTKLCCLPKETSKL   595
RGA3-blb   NL Q                 D S NFRI N   R  R          H  DS S    Q    595
SH10-tub   H EE S          M C D SEN  I                    HN YS S     P    597
RGA1-blb   E EQ    V           D S -NKIC   R                YN QS S    Q    604
B149-blb   E EQ    V           D S -NKIC   R    R           YN QS S    Q    604
SH20-tub   K EE                MD SNNIEI                      R       Q    594
T118-tar   K EE                MD SNNIEI                      R       Q    594
RGA4-blb   KLEQ         L      D SC-NNF    ER               VHN YS N    Q    648

Rpi-blb    GSLRNLLLDGSQSLTCMPPRIGSLTCLKTLGQFVVGRKKGYQLGELGNLNLYGSIKISHL   655
RGA3-blb           -C   ST     L     S SC  I KR        K          S TK   654
SH10-tub       FFH CDE NS         F    KWICC I       K RDV        E T   658
RGA1-blb   C    V  H-CP  S       L       Y    ER        R    R A S T   663
B149-blb   C    V DH CP  S       L       Y    ER        R    R A S T   663
SH20-tub        H CHR  RT                 S  K           S                654
T118-tar        H CHR  RT                                S                654
RGA4-blb   S  H VV  -CP  ST      L       F I  S          K    C   S T   707


Rpi-blb    ERVKNDKDAKEANLSAKGNLHSLSMSWNNFGPHIYESEEVKVLEALKPHSNLTSLKIYGF   715
RGA3-blb   D   K S           A     CL  DLD K R D ---E             KY E N   711
SH10-tub       VM                  I N SRKG          R I       P   C T S   716
RGA1-blb       ME          A         D--R  R                    P  KY E ID  721
B149-blb       ME          A         D--R NR                    P  KY E ID  719
SH20-tub        E          E       K DDDE  R       E              C   S    713
T118-tar        E          E       K DDDE  R       E              C T S    713
RGA4-blb        T  -       A  Q      D D  NR   K                  P  KY E IA  766

Rpi-blb    RGIHLPEWMNHSVLKNIVSILISNFRNCSCLPPFGDLPCLESLELHWGSADVEYVEEVDI   775
RGA3-blb   G  R  D   Q      V   R RGCE          E          T          DN--  769
SH10-tub     FRF            V   E  GCK          E    KR    QK   E     D----  774
RGA1-blb   C FC  D          V       GCE          E          QD  VE      DS--  779
B149-blb   C FC  D          V       GCE          E          QD  VE      D---  778
SH20-tub     R  D            L E   GCK                     YR   --            773
T118-tar     R  D            L E   GCK                   Q YR   --            773
RGA4-blb   G FRF S I      EKVI VR KSCK  L       E     N   QN   E       D--   824

Rpi-blb    DVHSGFPTRIRFPSLRKLDIWDFGSLKGLLKKEGEEQFPVLEEMIIHECPFLTLS-----   830
RGA3-blb   -  P -----          V    SN              K        TFYW  MFVIPTLSSV  823
SH10-tub   ---       R         F GE PN                    R T FY HMFVYTTL---  828
RGA1-blb   ----  L  R         H GG CN     QRMK  E          K SD  MFVFPTLSSV  835
B149-blb   ---   L  R         H GG CN     QRMK A           K SD  MFVFPTLSSV  832
SH20-tub     D         L      C CK DN          G           E RY  IP   -----   827
T118-tar     D                C CK DN     V    G           E RY  IP   -----   828
RGA4-blb       R S  RS    K  R  F R     M E    K  M     A LY  LFVFPTLSSV  884

Rpi-blb    -----------------SNLRALTSLRICYNKVATSFPEEMFKNLANLKYLTISRCNNLK   873
RGA3-blb   KTLKVI-ATDATVLRSI          D SN VE   L       S        N  FFR    882
SH10-tub   -----------------  F       H SH NE   L   I   SF       K  LFY    869
RGA1-blb   KKLEIWGEADAGGLSSI    ST    K FS HTV  LL        E  I   SV FLE    895
B149-blb   KKLEIWGEADAGGLSSI    ST    K FS HTV  LL        E  I   SV FLE    892
SH20-tub   -----------------P   K     N SD  E             S        N  HFK    868
T118-tar   -----------------    K     N SD  E           KS         N  HFK    869
RGA4-blb   KKLEVHGNTNTRGLSSI    ST      GA YR   L      TS T  EF SFFDFK     944

Rpi-blb    ELPTSLASLNALKSLKIQLCCALESLPEEGLEGLSSLTELFVEHCNMLKCLPEGLQHLTT   933
RGA3-blb                  FEF N           VK  T       S SN M              A  942
SH10-tub     S   C      T E HS S          VK  T         YD E   F          A  929
RGA1-blb             N  C D RY Y                                            955
B149-blb             N  C D RY Y                                            952
SH20-tub                    W      NI K  VK  T      I KFSKV         H   A  928
T118-tar                    W       I    VK  T      I KF K             A  929
```

```
RGA4-blb   D     T      R Q ES DS   F  Q     T   Q   KY K            A  1004

Rpi-blb    LTSLKIRGCPQLIKRCEKGIGEDWHKISHIPNVNIYI                        970
RGA3-blb    T T TQ  IVP    R         A   YLTL E                        979
SH10-tub      L R          ----------                                  948
RGA1-blb                                                               992
B149-blb                   ----------                                  971
SH20-tub    R   W          ----------                                  947
T118-tar    RV  W          ----------                                  948
RGA4-blb    N GVS   EVE    D E       A    LD H-                       1040
```

Figure 10B

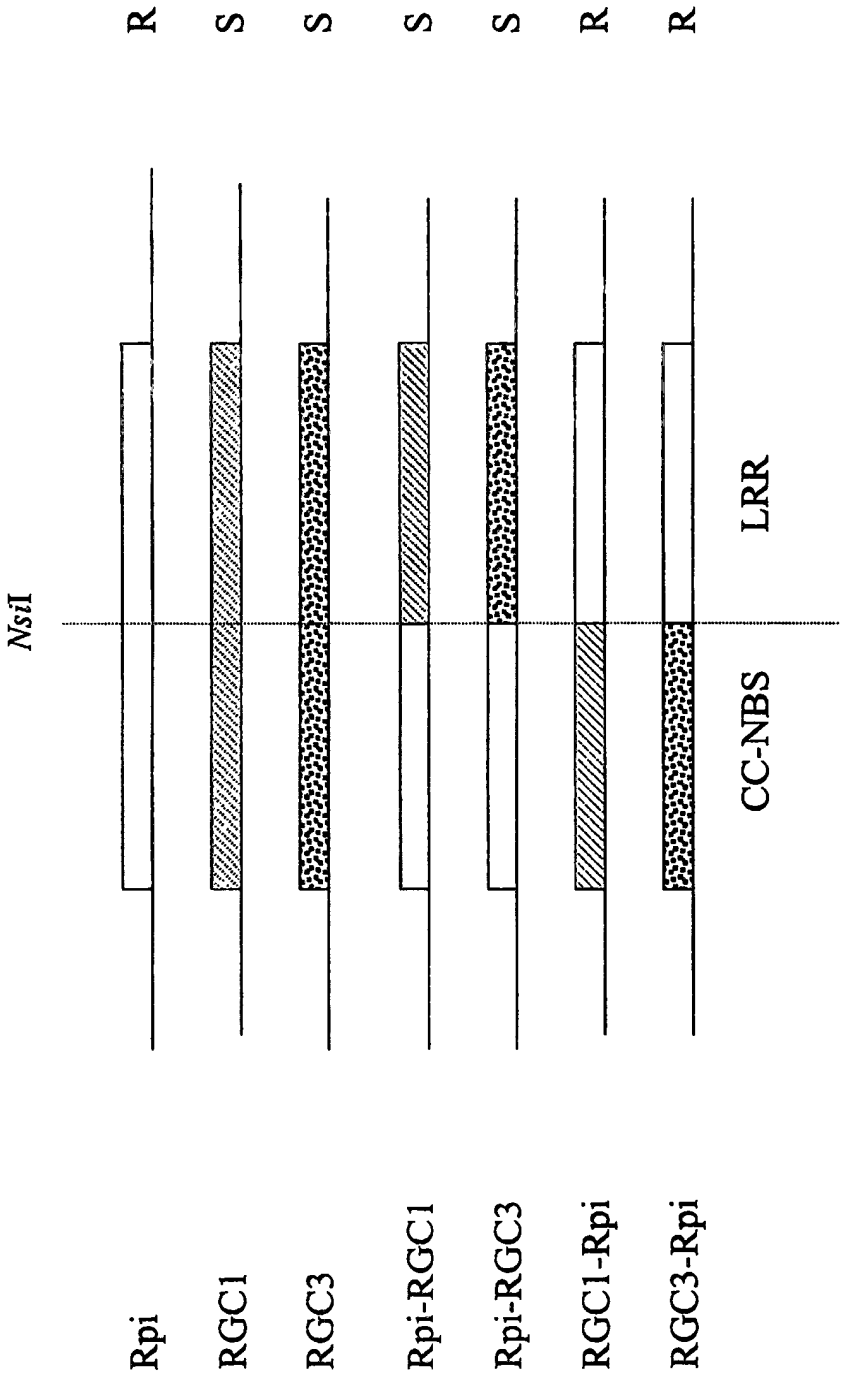
R
S
S
S
S
R
R
NstI
CC-NBS
LRR
Rpi
RGC1
RGC3
Rpi-RGC1
Rpi-RGC3
RGC1-Rpi
RGC3-Rpi

Figure 11

INTERNATIONAL SEARCH REPORT

Int. ional Application No
PCT/NL 03/00091

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 C07K14/415 C12N15/82 C07K16/16 G01N33/50 C12N5/10
A01H1/04

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 C12N C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)
EPO-Internal, BIOSIS, EMBASE, SEQUENCE SEARCH

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE EMBL 'Online! EMBL; 5 September 2001 (2001-09-05) PAN Q. ET AL.: "Lycopersicon esculentum isolate Q194 nucleotide binding region of resistance-like gene, partial sequence" Database accession no. AF404480 XP002206417	1-21
Y	abstract -& PAN Q. ET AL.: "Comparative genetics of nucleotide binding site-leucin rich repeat resistance gene homologs in the genomes of two dicotyledons: tomato and arabidopsis" GENETICS, vol. 155, no. 1, 2000, pages 309-322, XP002207023 --- -/--	22-29

[X] Further documents are listed in the continuation of box C. [ ] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 May 2003	01/07/2003

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Marinoni, J-C

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE SWISSPROT 'Online! EBI; 1 December 2001 (2001-12-01) SASAKI T. ET AL.: "Putative NBS-LRR type resistance protein" Database accession no. Q94J89 XP002206418	1-21
Y	abstract	22-29
X	DATABASE EMBL 'Online! EMBL; 8 June 2001 (2001-06-08) BOUGRI O. ET AL.: "Generations of ESTs from dormant potato tubers" Database accession no. BG890602 XP002206419 abstract	1-21
Y	VAN DER BIEZEN E A ET AL: "THE NB-ARC DOMAIN: A NOVEL SIGNALLING MOTIF SHARED BY PLANT RESISTANCE GENE PRODUCTS AND REGULATORS OF CELL DEATH IN ANIMALS" CURRENT BIOLOGY, CURRENT SCIENCE,, GB, vol. 8, no. 7, 26 March 1998 (1998-03-26), pages R226-R227, XP000924862 ISSN: 0960-9822 cited in the application the whole document	1-29
Y	LEISTER D ET AL: "A PCR-BASED APPROACH FOR ISOLATING PATHOGEN RESISTANCE GENES FROM POTATO WITH POTENTIAL FOR WIDE APPLICATON IN PLANTS" NATURE GENETICS, NEW YORK, NY, US, vol. 14, December 1996 (1996-12), pages 421-429, XP000964717 ISSN: 1061-4036 the whole document	1-29
A	VAN DER BIEZEN ERIC ET AL: "Plant disease-resistance proteins and the gene-for-gene concept" TIBS TRENDS IN BIOCHEMICAL SCIENCES, ELSEVIER PUBLICATION, CAMBRIDGE, EN, vol. 23, no. 12, December 1998 (1998-12), pages 454-456, XP002158209 ISSN: 0968-0004	
A	DONG F ET AL: "Development and applications of a set of chromosome-specific cytogenetic DNA markers in potato." THEORETICAL AND APPLIED GENETICS, vol. 101, no. 7, November 2000 (2000-11), pages 1001-1007, XP001087853 ISSN: 0040-5752	

-/--

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	ELLIS JEFF ET AL: "Structure, function and evolution of plant disease resistance genes." CURRENT OPINION IN PLANT BIOLOGY, vol. 3, no. 4, August 2000 (2000-08), pages 278-284, XP002206415 ISSN: 1369-5266	
A	YOUNG NEVIN DALE: "The genetic architecture of resistance." CURRENT OPINION IN PLANT BIOLOGY, vol. 3, no. 4, August 2000 (2000-08), pages 285-290, XP002206416 ISSN: 1369-5266	
A	OBERHAGEMANN P ET AL: "A GENETIC ANALYSIS OF QUANTITATIVE RESISTANCE TO LATE BLIGHT IN PATATO: TOWARDS MARKER-ASSISTED SELECTION" MOLECULAR BREEDING: NEW STRATEGIES IN PLANT IMPROVEMENT, KLUWER ACADEMIC PUBLISHERS, NL, vol. 5, no. 5, 1999, pages 399-415, XP001079515 ISSN: 1380-3743	
A	THIEME R ET AL: "PRODUCTION OF SOMATIC HYBRIDS BETWEEN S.TUBEROSUM L. AND LATE BLIGHT RESISTANT MEXICAN WILD POTATO SPECIES" EUPHYTICA, KLUWER ACADEMIC PRESS, AMSTERDAM, NL, vol. 97, no. 2, 1997, pages 189-200, XP002912898 ISSN: 0014-2336	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/NL 03/00091

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.: 30
because they relate to subject matter not required to be searched by this Authority, namely:

Rule 39.1(ii) PCT - Plant variety

2. [X] Claims Nos.: 1-30 all partially
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

see FURTHER INFORMATION sheet PCT/ISA/210

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 1-30 all partially

Claims 18 is directed to molecules binding either the nucleic acid of claims 1-7, whereas the application provides support and disclosure only for a limited number of such molecules i.e. primer/probes. The search has been restricted to primers/probes. All claims referring back directly or indirectly to claim 18 were partially searched too.

Additionally, claim 30 is directed to a plant that does not necessarily contain the gene(s) of the invention and could be a plant that has been obtained through traditional breeding methods, which are excluded from patentability under Article 53(b) EPC.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.